Illegitimate Recombination in *Escherichia coli*

by
Catherine E. Blake

Submitted for the degree of Doctor of Philosophy
Institute of Cell and Molecular Biology
University of Edinburgh
September 1996

Declaration

I hereby delcare that this thesis was composed by myself, and that the work described is my own unless otherwise stated.

Abstract

Illegitimate recombination is a powerful force that is capable of promoting large scale genome rearrangements without any requirement for the involvement of large regions of sequence homology. In this work an examination is made of palindrome instability, an example of an illegitimate event. The factors affecting deletion of long DNA palindromes from high copy number cloning vectors are investigated with particular reference to the mode of deletion and it is shown that the deletion of a 571 bp palindrome occurred from pMS7 using 3 bp repeats. A shorter 109 bp palindrome deleted from a related plasmid using 7 bp direct repeats and the mode of deletion is unaffected by the genotype of the host strain. There also appears to be a bias for the deletion of palindromic sequences on the lagging strand of a replication fork.

It is also shown that in a wild-type *E. coli* strain there is inhibition of plasmid multimerization if the plasmids carry long palindromic sequences. It is proposed that the lack of plasmid multimers in this background is a result of the removal of palindromic sequences form the plasmids by the SbcCD protein of *E. coli.* In an *sbcCD* strain, plasmid DNA bearing long palindromes is not detected in a monomeric form, instead the DNA is present in mutlimeric forms, predominantly dimers. The ability to form plasmid multimers in this background may help to

stabilise the palindromic sequences. The behaviour of palindromic sequences carried on plasmids is also investigated in *recA* and *recA sbcCD* strains with a view to the correct choice of *E. coli* strain for the cloning of long palindromic sequences.

Finally, the influence of an *sbcCD* mutation on the formation of *araB-lacZ* cistron fusion is investigated. The SbcCD proteins are thought to have a role in the processing of secondary structures formed by palindromic sequences. The formation of *araB-lacZ* fusions occurs via a strand transfer complex involving a complex genome rearrangment and secondary structure. Although an *sbcCD* mutation did not affect the kinetics or sequence specificity of fusion formation it is possible that SbcCD might have a role in processing the strand transfer complex which is not detected with the assay used.

Acknowlegdements

I wish to express my sincere gratitude to the following people:-

David Leach and Richard Hayward my supervisors who provided help and support wherever possible.

Noreen Murray for scientific advice and helpful gardening hints.

The rest of the group Thorsten, John, Angus, David, John, Lucy and Ewa along with the rest of the sixth workers.

To my drinking partners Thorsten, Pete, Laurie, Jim, Gareth, Ed, Steve, Jules, Dave and Richard who helped to put everything in perspective and made me laugh even when I did't want to.

To the wine producing nations of the world who I endeavoured to keep in profit during the period of my PhD, and the DSS and MRC who gave me the limited funds with which to do this.

To Isobel, Pauline and Aline who stopped me walking away from all of this when I didn't think I could go on anymore.

To many other friends including Rosalyn, Jenny, Helen and Ciara.

To my parents Edith and Richard for their never-ending supply of wine, Red Cross parcels and happiness. Also Andrew and Karen whose regular visits North injected a different kind of insanity into all of this and gave me an excuse for a break.

Table of Contents

Abbreviations

bp	base pair
kb	kilo base
kDa	kilo Dalton
in	inch
m	metre
lb	pound
mg	milligram
µg	microgram
ng	nanogram
λ	lambda
Δ	deleted

CHAPTER 1
INTRODUCTION

Introduction

In recent years there has been a rapid increase in the understanding of the proteins and mechanisms involved in homologous and site-specific recombination (for review see Kowalczykowski *et al.*, 1994; Weisberg & Landy, 1983). These processes are essential for the generation of genetic diversity, the repair of DNA and the correct segregation of chromosomes after replication. However, there is another equally powerful and often neglected mechanism capable of rearranging DNA. The term illegitimate recombination covers a class of recombination events that are not included in homologous or site-specific recombination. It was originally defined by Franklin (1971) as a recombination event between sequences of little or no homology and occurs independently of the *recA* function required for homologous recombination. In subsequent years, experiments by a variety of workers have done little to change this definition and illegitimate recombination has been demonstrated to be involved in deletion, duplication and amplification of DNA; transposon excision; cointegrate formation; the formation of specialised transducing phage and palindrome instability (for reviews see Weisberg and Adhya, 1977; Allgood and Silhavy, 1988). Illegitimate recombination has been detected in all organisms examined including phage, bacteria, yeast and humans (for reviews see Ehrlich, 1989 and Meuth, 1989). Although illegitimate events have been demonstrated to be independent of the proteins involved in either homologous or site-specific recombination, illegitimate recombination is still an example of a true recombination event as it results in the formation of a novel DNA joint i.e. the joining together of

two DNA segments that were previously non-adjacent. The lack of any requirement for sequence identity between two sequences involved in an illegitimate event means that potentially any part of the genome may recombine thereby making illegitimate recombination a major cause of genome alteration. Duplications provide additional copies of a gene that may accumulate mutations and so evolve new functions (Anderson and Roth, 1977). Deletions may alter the sequence context of a gene placing it under novel expression systems. Insertions of foreign genetic material into a genome may facilitate horizontal gene transfer so bypassing the need for similar gene functions to evolve repeatedly in different organisms (Arber, 1984). Illegitimate recombination events are also of great medical importance. They have been shown to be associated with or cause cancers such as human neoplasia (Croce, 1987), retinoblastoma and osteosarcoma (Friend *et al.*, 1986) along with a variety of human inherited disorders including thalassaemia (Collins and Weissman, 1984) and Duchenne muscular dystrophy (Forest *et al.*, 1987).

Illegitimate recombination is a broad term that covers too wide a range of subjects to be discussed here, as a result the remainder of this review will concentrate mainly on illegitimate recombination in *E. coli* discussing the factors involved and the proposed mechanisms for this mode of recombination. Particular attention will be paid to the involvement of illegitimate recombination in deletion formation and palindrome-mediated instability.

Genetic Rearrangements Caused by Illegitimate Recombination

Deletions

These are perhaps the most widely studied class of illegitimate events yet the actual details of deletion formation are still poorly understood, possibly because more than one mechanism may be involved (for review see Ehrlich *et al.*, 1993). Despite this apparent confusion there are several main features which define deletion formation.

RecA-independence

As mentioned previously illegitimate events occur independently of the RecA pathway for normal homologous recombination (Franklin, 1971; Jones *et al.*, 1982; Chedin *et al.;* 1994). However, this has not been seen to be true for every case studied. Albertini *et al.* (1982) found that deletion formation in their system occurred 25 times more frequently (per cell generation) in a *recA*$^+$ genetic background than in a *recA*$^-$ background. Control experiments argued that the result was not an artefact of the procedure used to determine mutation rates. A possible explanation for this result was proposed by Syvanen *et al.* (1986). Albertini *et al.* had selected *lacI* deletions in F'128 and it was later demonstrated that conjugal transfer during matings within F' populations led to elevated RecA activity. Syvanen

et al. (1986) suggested that the RecA-dependent stimulation that was observed by Albertini *et al.* was a consequence of the elevated RecA levels, which may have served to stabilise important intermediates in deletion formation. In this way the observed effect of RecA on the frequency of deletion formation may be related to conjugal transfer of F'128 and not the direct involvement of RecA in illegitimate recombination. Further support for this idea stems from the fact that the variety of deletions isolated by Albertini *et al.* in *recA* mutants was identical to that obtained in the wild-type background. Deletion formation has also been shown to be independent of other genes involved in the recombination and repair of DNA for example *recB, recC, uvrA, uvrB, uvrC* and *uvrD* (Anderson, 1970; Inselberg, 1967; Spudich *et al.*, 1970).

Sequence Homology

The lack of any requirement for long regions of sequence homology in deletion formation means that potentially any sequence may be involved in a deletion event, although any deletions that extend into essential sequences cannot be tolerated. Therefore, if deletions were formed totally at random the probability of isolating the same deletion more than once would be very small. However, some "hotspots" were identified that repeatedly led to the formation of identical deletions. Farabaugh *et al.* (1978) performed some of the earliest work on this subject. They examined nine novel joints formed by spontaneous deletions of 13-123 bp within the *lacI* gene of *E.*

coli which was carried on an F' plasmid. By comparing the nucleotide sequence of the novel joint with that of the wild-type sequence it was revealed that four of the nine deletions occurred at sites of microhomology consisting of 5-8 bp in direct repeat. Each of the deletion events had removed one of the direct repeats and the sequence in between. Three of the four deletions which exhibited microhomology were isolated twice, whereas the five deletions without microhomology were isolated only once.

Even more dramatic was the effect of microhomologies in the formation in large deletions. Albertini *et al.* (1982) analysed 24 deletions, ranging from 700-1000 bp, by direct DNA sequencing. Microhomologies of between 4-17 bp were used as deletion endpoints. Approximately 60% of the deletions isolated occurred between direct repeats of 17 bp in which 3 of the 17 bases were mismatched. Altering a single base within the homology decreased the deletion frequency by an order of magnitude, while introducing a base change into one of the 3 mismatched bases had no effect, thereby demonstrating the importance of homology length on deletion formation. The deletion frequency was also shown to be affected by the distance between the direct repeats with the frequency apparently decreasing with increasing distance. This observation was investigated in a more systematic manner by Chedin *et al.*,(1994). Deletion frequencies between 18 bp direct repeats were examined in a high copy number plasmid and the *B. subtilis* genome. In both cases deletion frequencies decreased exponentially by more than 1000-fold as the distance increased

from 33 to 2313 bp, indicating that the distance effect was not specific to the system used by Albertini *et al.* (1982). The decrease occurred in two phases with a transition at about 400bp.

Although a degree of sequence identity was present between the recombining sequences cited above, these events were unlikely to involve RecA-like functions since the lengths of the direct repeats were below the 30-70 bp required for homologous recombination in bacteria (Shen and Huang, 1986; Khasanov *et al.*, 1992). The lack of any obvious motifs in the short direct repeats makes unlikely that they would be recognised by enzymes which specifically cut and join DNA strands (Ehrlich *et al.*, 1993).

It should be noted that the presence of short sequence identities is not essential for deletion formation. In the examples discussed above, only half of those isolated by Farabaugh *et al.* (1978) used direct repeats and one deletion isolated by Albertini *et al.*. (1982) showed no identity between the recombining sequences. Benson and Bremer (1987) isolated deletions of 600-1000 bp that fused *lamB* to *lacZ*. Of the seven deletions detected only one was formed using limited microhomology (approximately 50% identity). Six other fusions displayed 20-35% homology in the direct repeats used. Sequence analysis demonstrated that other possible alignments with greater degrees of homology were present but these were never used as deletion endpoints. One interesting feature of the work by Benson and

Bremer is that the spectrum of deletions isolated was different when the selection was carried out on lactose minimal medium as opposed to lactose tetrazolium medium. This is consistent with the work of Shapiro (1984) on *lacZ* fusion formation which suggested that under non-lethal conditions medium composition, genetic background and the environment surrounding the potential clone all influence the selection process.

A model was proposed where deletion formation was mediated by palindromic or quasipalindromic sequences (Glickman and Ripley, 1984). It was suggested that these sequences could form secondary structures which could act as intermediates in deletion formation by bringing the direct repeats used as deletion endpoints closer together (the model will be discussed in more detail later in the chapter). Analysis of previously collected data (Farabaugh *et al.*, 1978) identified quasipalindromes at five of the nine deletion sites. Three of these also possessed direct repeats so that either or both could be involved in deletion formation. Deletions at the other two sites could involve quasipalindromes but not direct repeats. However, three of the nine deletions could not be explained by either model. Also, when the same model is applied to the deletions isolated by Benson and Bremer (1987), only two possessed quasipalindromes near the site of deletion formation. It is still unclear what is the major influence in the formation of deletions of the fused *lamB-lacZ* region. Trinh and Sinden (1993) cloned 17 and 18 bp palindromic and non-palindromic sequences into the *Eco*RI site of the CAT gene

present in pBR325. Deletion of the sequences occurred via the flanking direct repeats. A higher deletion frequency was observed for the palindromic sequence as compared to a non-palindromic sequence of the same length. Deletion of a non-palindromic sequence was influenced by the length of the flanking direct repeats.

Sequence Context

It is possible that more subtle features of chromosome sequence or structure may be involved in the cases where direct repeats or quasipalindromes cannot be identified at the sites of deletion formation. The chromosome is divided into approximately 40-50 topologically distinct domains (Sinden and Pettijohn, 1981). These domains may provide different environments for a particular sequence, with different levels of supercoiling or gene expression that could affect the frequency of deletion. This theory was tested by moving mutant β-lactamase genes from plasmid pBR322 to different sites within the *E. coli* chromosome (Kazic and Berg, 1990). The sequence to be deleted was made up of either a 22 bp or 90 bp palindrome flanked by three different 9 bp direct repeats and inserted into different sites in the *bla* gene along with the *lac* gene at 8' and a λ*lac* prophage at *att*λ and an F'*lac*. Moving the sequence from the multicopy plasmid to the single copy chromosome reduced the deletion frequency per cell by one or two orders of magnitude. This reduction was greater than could be accounted for by the reduction in copy number. The shorter palindrome deleted less frequently than the 90 bp palindrome. The

magnitude of the effect was also dependent on the position on the chromosome with a larger variation in reversion frequencies for integrated λ*lac* prophages than chromosomal *lac* constructs. Also some constructs which could be tolerated in λ*lac* could not be recovered as *lac* insertions. The two sites discussed lie in different chromosomal domains as proposed by (Rebollo *et al.*, 1988). It is possible that this accounts for the differences, although additional features may be involved.

Transcription induces changes in DNA topology; positive supercoils accumulate infront of the transcription complex and negative supercoils behind it (Pruss and Drlica, 1989). Vilette *et al.* (1992) used chimeric plasmids containing phage M13 and plasmid pBR322 sequences in *E. coli* to look at the effect of transcription on high frequency deletion events. In all the plasmids one deletion endpoint was located in the M13 replication origin nick site. The effect of transcription on the location of the other endpoint was examined by inserting promoters and terminators into the plasmids. Transcription affected the deletion in an orientation-dependent manner with more than 95% of deletion endpoints being located downstream of the induced promoter when it was in the same orientation as the major plasmid transcripts. The endpoints did not all lie within the transcribed region and they were not affected by the orientation of the pBR322 replication origin. It was proposed that deletion events occurred preferentially in a plasmid domain that was positively supercoiled by transcription and that this made the region more accessible to enzymes that were capable of introducing DNA breaks.

Duplications and Amplifications

Unlike deletions which are limited in size to non-essential regions of the genome, there is no upper limit on the size of duplications. Consequently duplications can arise at frequencies as high as 3% of the population (Anderson and Roth, 1981); this is much higher than the frequencies observed for the most active deletion hotspots. The only apparent constraint on the size of duplications is that they may not include the region for termination of replication (Anderson and Roth, 1981).

As with deletion formation the importance of short direct repeats in DNA duplications has been demonstrated. Edlund and Normark (1981) isolated a spontaneous duplication of 10 kb in the *ampC* region of the *E. coli* genome in which the recombination event occurred between 12 bp direct repeats. The role of microhomologies was confirmed by the work of Whoriskey *et al.,* (1987) who used an F' plasmid carrying a *lacI-lacZ* fusion with a *lacI* promoter deletion. *LacZ* duplications which fuse the *lacI-lacZ* hybrid to a distant active promoter were selected by growth on lactose. Investigation of the mutants isolated revealed that a single duplication did not produce a sufficiently high enough level of gene expression to allow growth on lactose. Instead, amplifications which increased the copy number of the duplication 40-200 fold were observed. Unlike the initial event, these further amplifications were RecA-dependent unequal crossover events.

Using sequence analysis Whoriskey *et al.* (1987) found that 28 of the 30 novel joints formed occurred between 5-18 bp direct repeats with one novel joint being present in 12 of the 30 duplications. As with deletion formation the presence of direct repeats was not essential for duplication since two of the novel joints produced lacked any sequence homology. It could be argued that the duplications isolated by Whoriskey *et al.* were not truly illegitimate events due to the involvement of RecA in the subsequent amplifications. However, the microhomologies used are below the limits needed for homologous recombination, so the initial duplication appear to be the result of illegitimate recombination. Given the similarities between duplication and deletion formation it seems likely that the two events occur by a related mechanism.

Cointegrate Formation

Unlike the genome rearrangements discussed above, cointegrate formation is an intermolecular event, but as with deletions and duplications the recombination can occur between short direct repeats. King *et al.* (1982) produced *in vitro* recombinants between phage λ and pBR322. Three of the four isolates resulted from a recombination event between 10 or 11 bp homologies. The fourth isolate was produced by two recombination events, both of which involved 10 and 13 bp homologies. Several other cointegrates were produced in a *recA* strain arose by recombination events between sequences showing homologies of less than 5 bp. The

recombinants were produced at a 10-fold lower frequency in a *recA* host. The structure of the recombinants also appeared to be different in a *recA* background (Marvo *et al.*, 1983). Of the four cointegrates sequenced, all had deleted DNA around the fusion joint. One of the cointegrates had also duplicated part of the plasmid sequence during the fusion event. Owing to the small numbers of cointegrates examined in the study it is unclear whether the differences observed in *recA*$^+$ and *recA*$^-$ hosts are due to different mechanisms being active in each background.

The effects of mutations in other genes involved in DNA recombination and repair were also investigated. Ikeda *at al* (1980) added plasmid DNA to a λDNA *in vitro* packaging system and isolated phage which carried plasmid-borne resistance markers. The frequency was not altered by *int* or *red* mutations in the λ prophages or by *recA, recB recC* or *recF* mutations in the *E. coli* host from which the packaging extracts were isolated.

Formation of Specialised Transducing Phage

At its most basic the formation of specialised transducing phage can be thought of as a deletion event which results in the formation of a covalently closed circular molecule composed of the deleted sequence. If the DNA molecule contains an origin of replication it may be recovered. Unlike general sequence deletions by

illegitimate recombination, the formation of specialised transducing phage were characterised by their large size, typically ~50 kb. Specialised transducing phage are produced by aberrant excision events which results in bacterial genes adjacent to the prophages becoming fused to the phage genome. To be recovered the excised phage must be of a suitable length to be packageable and for the case of λ phage have cohesive ends which are still intact (Kaiser and Wu, 1968). Lysates of induced λ contain phage able to transduce *gal* and *bio* genes to the host cell at a frequency of ~ 10^{-6} (Campbell, 1964). Of the various *E. coli* gene functions affecting DNA recombination and repair, many have found to be non-essential for the excision of λ transducing phage. Mutations in *recA* and *recB* do not affect the yields of specialised transducing phage (Franklin, 1967). Mutations affecting the repair of U.V. damage (*uvrA, B, C,* or *D*) cause a 5-fold or less decrease in *bio* transducing titres of λ lysates (Franklin, 1967). Recent work by Ukita and Ikeda (1996) identified a recombination hotspot exhibiting short regions of homology which accounted for 60% of all *λbio* phage produced in a wild-type *E. coli.* This hotspot was not detected in a *recJ* host.

Transposon Excision

When most transposons insert into a genome the target sequence is duplicated, resulting in the inserted transposon being flanked by the direct repeats of the target sequence. Recombination between the direct repeats of the duplicated

target sequence lead to the excision of the transposon and the retention of one copy of the duplicated sequence (Kleckner, 1979; Berg *et al.*, 1980).

Transposon Tn5 is 5.7 kb in length, contains 1.5 kb inverted terminal repeats (IS50 elements) and has a central region which codes for kanamycin resistance (Berg, 1989). Tn5 duplicates a 9 bp target sequence on insertion. Transposase, the protein required for transposition of Tn5, is not required for excision of the transposon. As with the deletion of others sequences, the frequency of excision of Tn5 varies greatly depending on the insertion position of the transposon. It ranged from $3x10^{-8}$ to $2x10^{-4}$ for different positions within the *lac* gene of an F'*lac* (Berg *et al.*, 1980). When a Tn5 which has the IS50 elements in direct repeat as opposed to the normal inverted repeat, the excision frequency is reduced from 10^{-6} to 10^{-9} suggesting a role for the inverted repeats in the excision event (Egner and Berg, 1981).

Excision of transposon Tn10 is similar to that of Tn5. Tn10 is 9.3 kb in length and encodes tetracycline resistance (Kleckner, 1989). It duplicated a 9 bp target sequence on insertion and can excise by recombination between the direct repeats (Kleckner, 1979). In addition, Tn10 can undergo nearly precise excision by recombination between 23-24bp direct repeats present in the 1.4 kb inverted repeats of the flanking IS10 elements. Nearly precise excision leaves 50 bp of the transposon flanked by the 9 bp target repeats in the host genome. This sequence can

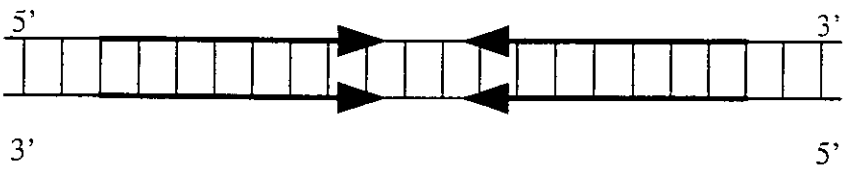
then undergo deletion via the 9 bp direct repeats (Foster *et al.*, 1981). All three examples of excision in Tn10 occur independently of the transposase function. Excision does not lead to transposition to a new site and occurs independently of host recombination functions such as *recA* (Foster *et al.*, 1981).

There is evidence that transposon excision occurs more frequently from replicons using rolling circle replication as opposed to theta replication (Berg *et al.*, 1983). A chimera was made between the ColEI replicon and the M13 origin and activation of the M13 origin induces rolling circle replication. The plasmid carried an ampicillin resistance gene inactivated by transposon insertion. Nearly precise excision of the transposon restores ampicillin resistance. Infection of cells bearing the chimeric plasmid with the helper phage fI led to a 1000-fold increase in ampicillin resistance cells within three hours. It is possible that the single stranded nature of rolling circle replication allows the alignment of direct repeats used in deletion formation.

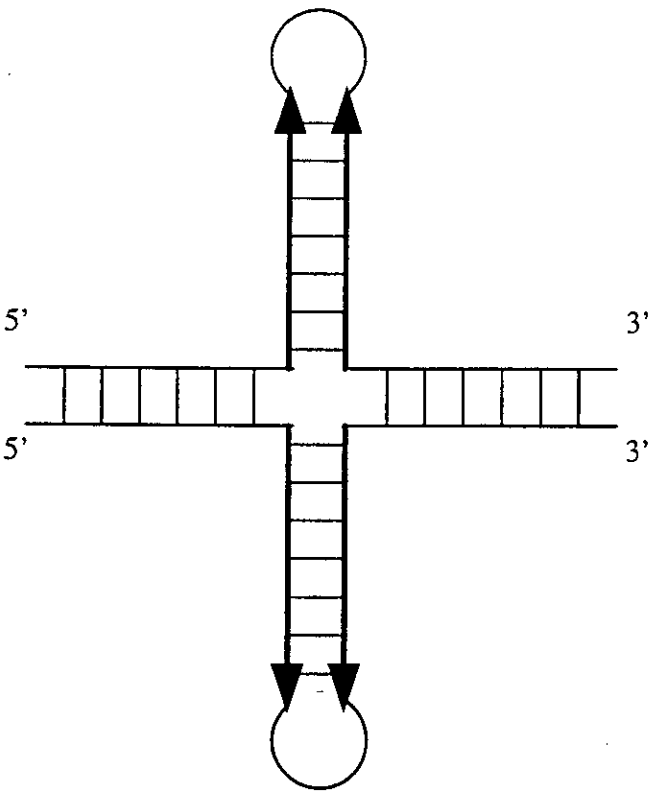
Palindromic DNA Sequences

A palindrome is an inverted repeat sequence of DNA with two-fold rotational symmetry. Its sequence arrangement allows it to exist in two forms; either an interstrand base paired linear form or a cruciform structure with intrastrand base

Figure 1. Secondary structure of palindromic DNA.



1a.



1b.

1a. Palindrome in linear interstrand base paired conformation.

1b. Intrastrand base paired to form a cruciform structure.

⟶ Inverted repeat sequence.

pairing (fig. 1). Long perfect palindromes (>40 bp in length) do not naturally occur in the *E. coli* genome. However, small (<40 bp) imperfect palindromes are found in extragenic sequences e.g. REPs (Young and Ames, 1990), BIMEs (Bachellier *et al.*, 1994) and IRUs (Sharples and Lloyd, 1990). Imperfect palindromes are also located in genetic regulatory regions, for example, in the *lac* operator, CAP protein binding site (Bertrand *et al.*, 1975) and in the initiation of DNA replication (Hirota *et al.*, 1979). In the early days of DNA cloning experiments it was observed that although directly repeated cloned fragments were occasionally recovered in ligations, invertedly repeated sequences were never recovered in transformed bacteria (Behnke *et al.*, 1979; Casadaban and Cohen, 1980). This prompted the suggestion that such palindromic sequences could not be tolerated in *E. coli* perhaps because of their ability to adopt unusual secondary structures. When long perfect palindromes, either obtained from eukaryotic cells where they are prevalent or constructed *in vitro*, were introduced into a bacterial cell they exhibited the two distinct effects of inviability and instability.

Inviability

This results from the loss of the replicon carrying the palindrome and there is an upper limit of approximately 150-200 bp in total length before a palindrome exhibits inviability (Warren and Green, 1985). Replicon inviability can be overcome by deletion of the entire palindromic sequence, deletion of the central part of the

palindrome or insertion of segments of DNA into the centre of the palindrome to separate the inverted repeats. Inviability is also strongly influenced by the genotype of the host used. Leach and Stahl (1983) reported that a lambda phage bearing a 3.2 kb perfect palindrome plated at an efficiency of less than 1% of the control phage lacking the palindrome when in a *rec*$^+$ host. However, *sbcC* is a gene which when mutated greatly alleviates inviability. Originally characterised as an additional mutation in *recBC sbcB* strains (Lloyd and Buckman, 1985), mutation of the single *sbcC* gene was sufficient to overcome inviability and allow replication of a 571 bp nearly perfect palindrome in *E. coli* (Chalker *et al.*, 1988). SbcC along with SbcD form the primary control for the replication of long palindromes in *E. coli*. The DNA sequence of the genes revealed that the gene products were distantly related to the exonucleases of bacteriophages T4 and T5 responsible for the degradation of chromosomal DNA (Leach *et al.*, 1992). Genetic evidence suggested that mutation of the *sbcC* gene may mimic the action of the *gam* gene of bacteriophage λ (Kulkarni and Stahl, 1989) by allowing viability of long palindromic sequences.

These two lines of evidence suggested that the Gam protein might interact with the SbcC protein. Since the Gam protein is known to inactivate the RecBCD enzyme (Karu, 1975) it was suggested that Gam may also inactivate the SbcC protein. The gene products of the *sbcCD* genes have been purified and shown to possess an ATP-dependent double-strand exonuclease activity (Connelly and Leach, 1995). *SbcCD* mutations were originally isolated as additional mutations that

increased the viability of *recBC sbcB* strains. Plasmids propagated in these strains also give rise to high levels of linear multimeric DNA possibly because of uncontrolled rolling circle replication. The activity of the SbcCD proteins could be explained in these two examples by digestion of the linear substrate for recombination in the first example and by the digestion of rolling circle tails in the second example. The identification of the *sbcCD* genes has allowed the study of palindromic sequences *in vivo*. Although mutation of *sbcC* overcomes the primary determinant of inviability, palindromes retain some problems in this background. This is reflected in the reduced plaque size for lambda phage containing long palindromes. Also, instability is still seen with palindromes which are no longer inviable in this background.

Instability

This results in deletions which partially or wholly remove the palindromic sequence at a high frequency (Collins, 1981) and many of the features of general sequence deletions apply to palindrome instability. Unlike inviability, instability has been detected for palindromes as short as 22 bp in length (DasGupta *et al.*, 1987). Palindrome-mediated deletions, like spontaneous deletions in general sequence DNA, occur by a *recA*-independent mechanism. Generally, the level of instability increases with length of the palindromic sequence (Weston-Hafer and Berg, 1989). The deletion frequency is also affected by the general sequence context as well as the

local DNA context, such as the presence of flanking direct repeats (Farabaugh *et al.*, 1978; Albertini *et al.*, 1982), the very nature of cloning palindromic sequences means that they are flanked by directly repeated restriction sites.

The overall sequence context of the palindrome is important. For example, it was demonstrated by DasGupta *et al.* (1987) that the deletion frequency of 22, 32 and 90 bp palindromes, derived from Tn5, varied 100-fold for the same palindrome at different insertion sites. A 3000-fold difference in deletion frequency was observed for a 90 bp palindrome at a pair of adjacent sites which overlapped at 8 out of the 9 bp. This suggests that differences in deletion frequency must involve precise effects of local DNA sequence. Excision of the 22 bp palindrome was 3-550 times more efficient, than for excision of the whole 50 kb transposon from which it was derived, at a number of different sites. For all the insertions (except one) a small increase in deletion frequency was observed when the length of the palindrome was increased to 32 bp. Increasing the length of the palindrome to 90 bp increased the frequency of deletion 9-314 times indicating an increase in instability with increasing length of inverted repeats. The one insert which did exhibit a significant increase in deletion frequency had a deletion frequency of 116 and 18000 times when the length was increased from 22 bp to 32 bp and 90 bp respectively at the same site.

Kazic and Berg (1990) placed a variety of deletion-prone sequences into the same immediate context i.e. within a *lacZ* gene present at different locations within

genomes. The positions used were at λ*lac* at 17' on the *E. coli* genome, chromosomal *lac* at 8' and on an F'*lac*. Palindromes which were 22 and 90 bp in length, flanked by 4 bp direct repeats, were placed in the *bla* gene of pBR322 and the gene moved to the different *lac* locations. Deletion of the palindrome resulted in a reversion to ampicillin sensitivity. The 22 bp palindrome deleted less frequently than the 90 bp palindrome (~10^{-2} fold less). When moved to chromosomal locations most of the constructs exhibited a reduced level of deletion (1-2 orders of magnitude). The magnitude of the effect was insertion-specific and location-specific. This reduction was beyond any effect due to the drop in copy number which resulted in moving the constructs from a multicopy plasmid to single copy genomes. All *lac* positions were not identical in their effect on the deletion frequency since some constructs which were tolerated at the λ*lac* position could not be recovered from the chromosomal *lac* position.

Both DasGupta *et al.* (1987) and Kazic and Berg (1990) demonstrated the increase in deletion frequency when the length of the palindromic sequence is increased from 22 to 90 bp. Weston-Hafer and Berg (1989) used a variety of palindromes to look at the effect of palindrome length on deletion from identical positions within pBR322 and also the choice of direct repeats used as deletion endpoints in the recombination reaction. 19, 22 and 26 bp palindromes showed a higher deletion frequency when flanked by 10 bp repeats as opposed to 9 bp repeats. 4 bp direct repeats were also present (one in the palindrome and one in the plasmid)

but for the 22 bp palindrome these repeats were used in only 2% of deletion events. Palindromes with increased lengths (32, 34, 48 and 90 bp) exhibited increased deletion frequency when flanked by 9bp repeats rather than 10 bp repeats. With palindromes which were 36, 44 and 90 bp in length 80% of the deletion events used the 4 bp direct repeats. Use of these shorter repeats increased form 2-56% when the length of the palindrome was increased from 22 to 32 bp. It was suggested that the use of shorter direct repeats reflected the increased stability of secondary structures formed by the palindromes. Replication may be able to proceed into a longer palindrome as far as the internal 4 bp repeat without disrupting the structure, whereas with a smaller palindrome the replication machinery would disrupt the secondary structure and so be able to successfully replicate the sequence.

The presence of any central asymmetric region in the palindrome decreases the deletion frequency and, if long enough, may confer stability to palindromes (Warren and Green, 1985). Plasmid pAT153 was restricted with *Sau*3A to give a large number of restriction fragments which varied in length from 8-876 bp and the fragments were ligated into the centre of a long palindromic sequence. The smallest insertion recovered was 72 bp made up of the insertion of a 46 bp and a 36 bp fragment suggesting that insertion of less than about 50 bp did not alleviate inviablity. The palindrome remained highly unstable and suffered deletions. Instability was detected in all constructs which carried an insertion of less than 150 bp. A more natural example of the effect of central insertion is seen with transposons

which are in essence made up of long inverted repeats separated by a large central insertion (Berg, 1989; Kleckner,1989).

Genes Involved in Illegitimate Recombination

Many of the studies investigating the mechanisms of illegitimate recombination have relied on screening for chromosomal loci which can be mutated to increase the frequency of deletion formation. Such mutations could identify genes whose products play a role in illegitimate events.

Initial screening looked at genes which were involved in DNA recombination and repair. Mutations in *recA* , *recB, recC, uvrA, uvrB, uvrC* and *uvrD* were all found to have no effect (Franklin, 1967; Inselberg, 1967; Spudich *et al.,* 1970; Coukell and Yanofsky, 1970). Mutations isolated in *polA* were shown to increase the deletion frequency in the *tonB-trp* region of the *E. coli* chromosome (Coukell and Yanofsky, 1970). However, the allele used was usually *polAI* which causes a dramatic decrease in polymerase function with little effect on the exonuclease activity of the enzyme (Joyce *et al.,* 1985). It seemed unlikely that *polA* mutations were having a direct involvement in illegitimate recombination, but may have given rise to an increase in gapped and damaged DNA which may have acted as a substrate for illegitimate events.

Ligase mutations were isolated that gave rise to increased deletion frequencies in the *tonB-trp* region of the *E. coli* genome (Gellert and Bullock, 1970). Subsequently work by Shafferman *et al.* (1987) demonstrated an increase in palindrome deletion from a plasmid in a *lig*^{ts} mutant at the non-permissive temperature. However, it was proposed that the lack of a functional ligase might lead to an increase in the presence of gapped DNA within the cell which may be prone to deletion.

A number of mutants (called *tex* mutants) have been isolated which exhibit increased excision of transposable elements. Many of the mutations mapped to previously identified genes involved in mismatch repair e.g. *dam, mutH, mutL, mutS mutU* and also to *recB* and *recC* (Lundblad and Kleckner, 1982, 1985; Lundblad *et al.*, 1984). The highest frequencies of deletion were seen with mutations in *mutU* and *recBC* (~100 fold), the lowest with *mutH* and *mutS* (~2 fold). However the same mutations did not affect excision from different sites to the same degree (particularly for mutations in *mutS*). Whether the effect of these mutations was a direct one was not clear but the effect of the mismatch repair mutations was decreased by decreasing the size of the transposon inverted repeats (Lundblad and Kleckner, 1985). A mutation isolated in *dnaQ* (*mutD*), which encodes the proof-reading subunit of polymeraseIII (Scheuermann *et al.*, 1983), gave rise to increased levels of base mismatches (Lundblad and Kleckner, 1985) which may also stimulate an illegitimate event rather than there being a direct involvement of DnaQ.

Lejeune and Danchin (1990) reported that deletion of the *bglY* gene in *E. coli* resulted in an increased frequency of large deletions of chromosomal markers and plasmid DNA. The frequency of point mutations was unaffected. It was proposed that the *bglY* gene product, a small histone-like protein (Hulton *et al.*, 1990), could be a factor involved in the control of tertiary DNA structure. The absence of the protein could lead to the looping out of single stranded DNA so increasing the chances of short direct repeats aligning and subsequent deletion of the intervening sequence.

It was also reported that an *E. coli* strain carrying a mutation in the *xth-pnc* region at 38' on the chromosome again showed an increased frequency of illegitimate recombination (Yi *et al.*, 1988). Subsequently Whoriskey *et al.* (1991) identified a gene, *mutR,* which mapped near 38.5' on the *E. coli* chromosome. The mutation resulted in increased deletion formation on multicopy plasmids and it seemed likely that removal of *mutR* was responsible for the phenotype demonstrated by Yi *et al.* (1988). Cloning and sequencing of *mutR* (Schofield *et al.*, 1992) revealed that the gene was identical to *topB* a gene which encodes topoisomerase III (DiGate and Marians, 1988). Topoisomerase III has been shown to relax negatively supercoiled plasmids so the lack of topoisomerase III might lead to an increase in supercoiled structures which might increase deletion formation (Schofield *et al.*, 1992). However, the primary activity of topoisomerase III seems to be decatenation so how this would affect deletion formation is unknown.

Mutations which mapped to the *sbcB* gene at 43.5' on the *E. coli* chromosome resulted in increased deletion formation on a plasmid that created rare *lacZ* fusion proteins (Allgood and Silhavy, 1991). The only known biochemical activity of the *sbcB* gene product is that of a single-stranded 3'-5' processive exonuclease (Kushner *et al.*, 1971). The mutant alleles isolated were defective in this function. As a result it was suggested that Philips *et al.* (1988) were correct in indicating that the SbcB protein had a second, unidentified function which might play a role in illegitimate recombination.

The formation of specialised transducing phage in *E.coli* is a rare event but it can be induced by U.V. irradiation. A mutation in the*E. coli recJ* gene was found to decrease the formation of λ*bio* specialised transducing phage by 3-10 fold (Ukita and Ikeda, 1996). A recombination hotspot which gave rise to approximately 60% of all λ*bio* phage produced in a wild-type *E. coli* was not detected in a *recJ* strain. It was proposed that RecJ preferentially promoted illegitimate recombination at that hotspot. Both this hotspot and other sites of recombination which gave rise to transducing phage exhibited short regions of homology (3-10 bp). All recombination events at the major hotspot used 9 bp direct repeats. A model for recombination was proposed in which a double-strand break induced by direct repeats was acted on by the 5'-3' exonuclease activity of RecJ (Yamaguchi *et al.*, 1995). In the model a lesion induced by u.v. irradiation blocked the progression of a replication fork leading to slippage between the direct repeats and the formation of a DNA loop.

Unwinding of the DNA would produce single-stranded DNA subject to attack by RecJ which is known to possess a single-stranded DNA 5'-3' exonuclease activity (Lovett and Kolodner, 1991). Joining of the DNA could then take place using the complementary direct repeats.

Extensive work has been carried out to identify a role for DNA topoisomerases in illegitimate recombination. DNA topoisomerases are found in bacteria (Wang, 1991) as well as the nuclei and mitochondria of eukaryotes (Champoux and Dulbecco 1972; Fairfield *et al.*, 1979) and can be classified into two categories. Type I make single strand DNA breaks and relax DNA supercoiling by changing the linking number in steps of one. In *E. coli* one such enzyme is topoisomerase I (omega protein) which is a 101kDa polypeptide encoded by the *topA* gene. Type II topoisomerases make double strand breaks and mediate the interconversion of supercoiled and relaxed forms of DNA by changing the linking number in steps of two. In *E. coli* an example of a type II enzyme is DNA gyrase. It is composed of two subunits of 105kDa and 95kDa encoded by the *gyrA* and *gyrB* genes respectively and has a tetrameric structure of A_2B_2 (Gellert, 1981; Drlica, 1984).

Ikeda *et al.* (1981; 1982 and 1984) developed an *in vitro* system which demonstrated that DNA gyrase could promote illegitimate recombination in *E. coli*. The incubation of plasmid pBR322 with phage λ in *E. coli* extracts led to the

formation of lambda-pBR322 recombinants. The recombinants were formed by the insertion of the plasmid into λ DNA or by the substitution of lambda DNA by plasmid DNA. The sites of crossovers appeared to be random implying the involvement of illegitimate recombination. The reaction was found to occur independently of the *E. coli* RecA function as well as the *int* and *red* gene products. The addition of oxolinic acid, a DNA gyrase inhibitor which acts on the A subunit stabilising the enzyme/DNA complex (Gellert *et al.*, 1977), stimulated recombination 13 fold. The effect was specifically on gyrase since the use of an extract from a *gyrA* mutant did not produce recombinants.

The structure of recombinants between λ and pBR322 formed in the presence of oxolinic acid were analysed by heteroduplex analysis and DNA sequencing (Ikeda *et al.*, 1982; Naito *et al.*, 1984). Among nine isolates investigated, two were formed by the direct insertion of the plasmid into λ. In other cases the phage or the plasmid suffered deletions upon insertion. In all cases where a deletion occurred one end of the deletion coincided with one end of the plasmid insertion point. Recombination sites were randomly distributed on the phage and plasmid genomes. Sequence comparisons revealed that the sites did not display more than 4 bp homology and in one of the recombinants no homology was present. It was proposed that homology at deletion endpoints was not necessary for gyrase-mediated illegitimate recombination.

It was noted by Marvo *et al.*, (1983) that some of the recombination sites resembled those of the gyrase cleavage consensus. The cleavage sequences for gyrase have been determined *in vitro* as 5' YRT ↑ GNYNNY 3' (Morrison and Cozarelli, 1979) and *in vivo* as 5' NRT ↑ GRY CT/CY 3' (Lockshon and Morris, 1985). The consensus sequence obtained for the phage/plasmid recombinants was 5'NRT ↑ RNNYNY 3' which only weakly matches the consensus cleavage sites for gyrase. Also the *in vitro* and *in vivo* cleavage sites were determined using pBR322 DNA and none of these cleavage sites were used in the formation of the λ/plasmid recombinants.

Direct tests were also carried out *in vivo* (Ikeda, 1990) to see if DNA gyrase participated in illegitimate recombination in *E. coli* by isolating *gyrA* mutations which reduced the deletion frequency of pBR322 from λ-pBR322 hybrids. In contrast, oxolinic acid increased the excision rate of pBR322 approximately sixfold (Ikeda, 1990). Similarly, a $gyrA^{ts}$ mutation was found to stabilise a 2.1kb palindrome isolated from a slime mould, cloned into plasmid pAG60 (Saing *et al.* , 1988).

From this data Ikeda *et al.* (1982) proposed a model for DNA gyrase-mediated illegitimate recombination (figure 2). It was thought that gyrase binds to DNA as a A_2B_2 complex and cleaves the double strands of the DNA resulting in an intermediate in which each gyrA subunit covalently binds the 5' end of the DNA at

the cleavage site (figure 2b). Two gyrase/DNA complexes then form a tetramer (A_4B_4) (figure 2c). Dissociation of the tetramer into two dimers might lead to subunit exchange and so an exchange of DNA strands (figure 2d). Although this mechanism may act *in vivo* to promote illegitimate recombination it seems likely that this is not responsible for all known illegitimate recombination events.

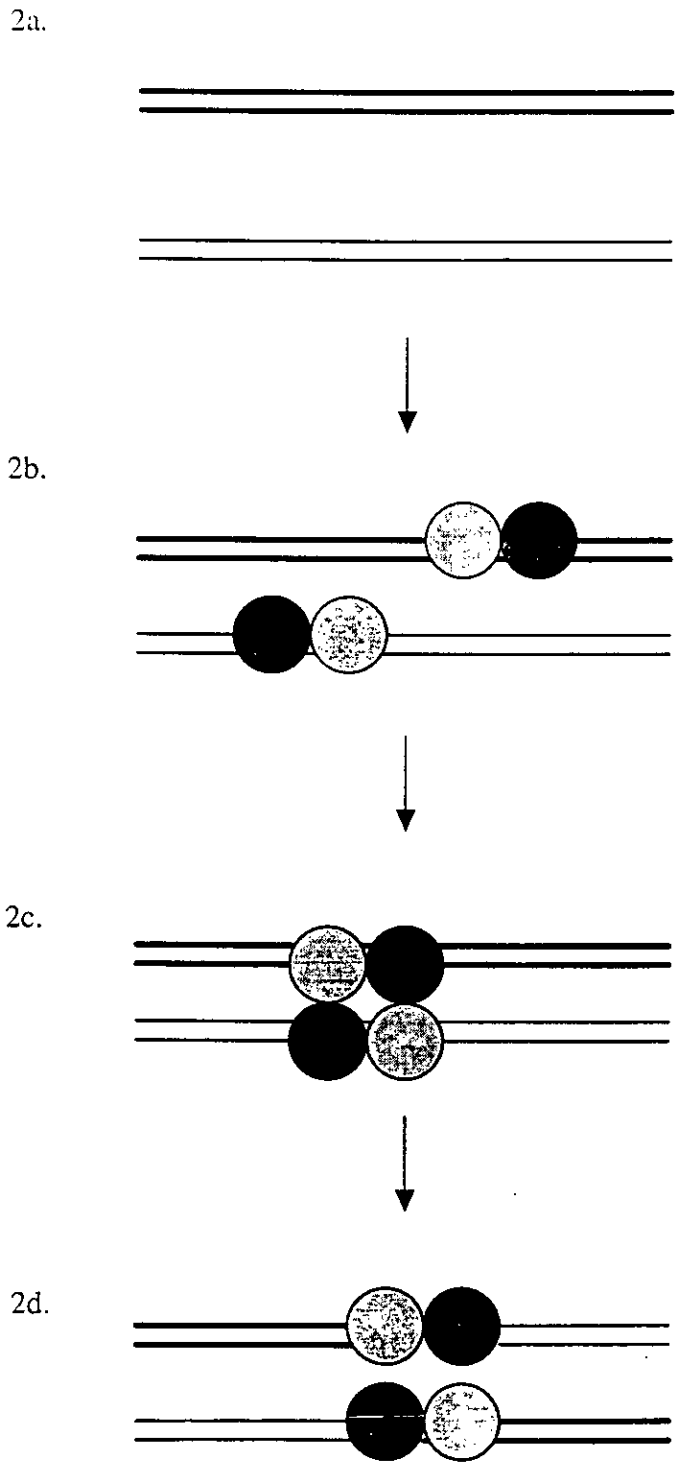
Possible Mechanisms for Illegitimate Recombination

Two fundamentally different mechanisms for illegitimate recombination, based on the available evidence, have been proposed. The first model, termed replication slippage, stems from work by Streisinger *et al.*, (1967) on the mechanism of frameshift mutations. This model has been adapted to account for palindrome-stimulated events. In this model (fig.3), the polymerase stalls due to the presence of a stem-loop structure in the DNA template caused by intrastrand base pairing in palindromic DNA, the presence of quasipalindromic sequence or single stranded loop

Figure 2

Proposed model for DNA gyrase mediated illegitimate recombination. Gyrase binds to DNA molecules as A_2B_2 complex at cleavage site (2b). The two gyrase/DNA complexes form a tetramer (2c). Subunit and associated strand exchange occurs followed by dissociation of the teramer (2d).

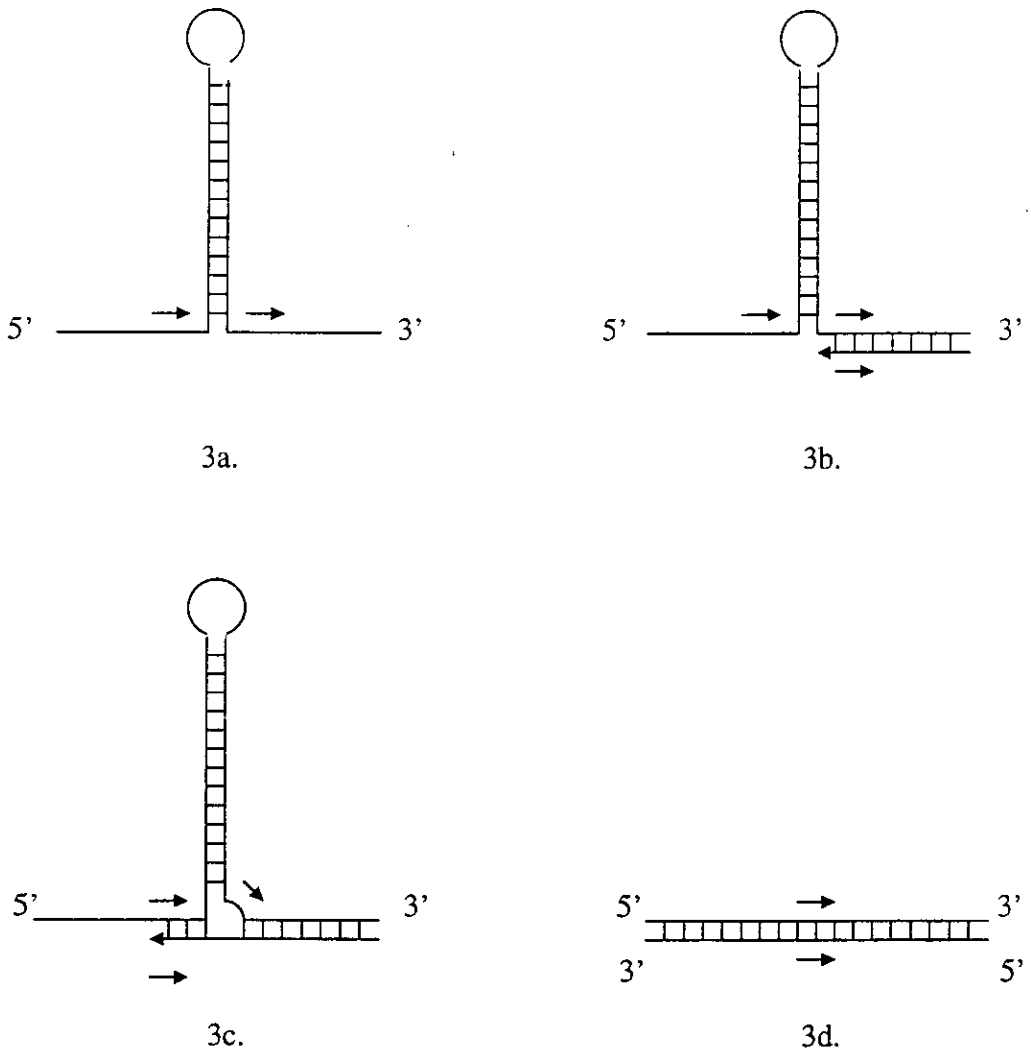
Figure 2. Model for gyrase-mediated illegitimate recombination (after Ikeda *et al.*, 1990)



in general sequence DNA or cruciform extrusion. The presence of a structure would increase the chances of the nascent DNA strand dissociating from the template strand and reannealing to a second copy of a directly repeated sequence on the far side of the structure. Palindromes flanked by short direct repeats, such as restriction target sites, seldom use these repeats for deletion (Kazic and Berg, 1990; Weston-Hafer and Berg, 1991). Instead other direct repeats are used, with a preference for one of the repeats to be located just within the palindrome and one to be located just downstream. This is consistent with the replication progressing some way into the structure before dissociation of the nascent strand, which could then anneal to the repeat copy downstream. Replication of the DNA would then continue, the result being that one of the repeat copies has been deleted along with the intervening sequence. This model could also be used to explain duplications and amplifications, which would be created if replication slippage resulted in the sequence between the direct repeats being repeatedly copied. This could be a result of slippage of the second copy of the repeat in the daughter strand to the first copy of the repeat in the template strand.

The asymmetry of the deletion event permits the question to be asked of whether there is a bias for the deletion to occur on the lagging or leading strands of the replication fork. The direct repeats can be defined as donor and target, with the donor sequence being in the palindrome and the target located downstream. Inversion of a DNA sequence containing a palindrome and the favoured target site moves the

Figure 3. Replication slippage model for deletion formation (after DasGupta *et al.*, 1987).



3a. Hairpin structure formed during DNA replication.

3b. Polymerase stalls at the hairpin.

3c. Dissociation of nascent strand, followed by re-annealing to a direct repeat sequence on the far side of the hairpin.

3d. Re-initiation of replication with subsequent loss of one direct repeat and the intervening sequence.

preferred orientation of repeats from leading to lagging strand with respect to the direction of replication. Work by Trinh and Sinden (1991) suggested that deletion occurred preferentially on the lagging strand. However, experiments performed by Weston-Hafer and Berg (1991) showed no difference in deletion frequency between the two strands.

Evidence exists to suggest that polymerases will stall at hairpin structures in the template *in vitro*. LaDuca *et al.*(1983) found sites on an fd phage DNA template which blocked the progression of four forms of *E. coli* DNA polymerase III. The progression of the enzyme was found to be highly sensitive to any secondary structure present in the template strand and to the base composition of the DNA. Approximately 65% of pause sites were within 15 nucleotides of potential hairpin structures. The remainder of pause sites occurred where there was no discernible secondary structure. Hairpin structures with high GC base composition in the stem acted as the strongest pause sites. The presence of additional subunits associated with the core enzyme altered recognition of some of the pause sites and coating of the template strand with single-stranded binding protein (SSB) reduced the level of polymerase stalling (Sherman and Gefter, 1976).

Not all pause sites could be correlated with the presence of DNA able to adopt a secondary structure. *In vitro* studies performed by Weaver and DePamphilis (1984) again identified a number of sites in M13 derivatives which could arrest the

movement of the mammalian DNA polymerase α. They defined two classes of pause sites: class I could be correlated witht the presence of palindromic DNA. Class II sites exhibited no secondary structure but they did contain either direct repeats between 8-22 bases in length, or alternating purine and pyrimidine residues. A stretch of such residues strongly arrested polymerase α as it had just completed their synthesis.

Evidence exists in favour of replication slippage in illegitimate recombination, mediated by direct repeats. d'Alencon *et al.* (1994) looked at transposon excision from M13/ColE1 chimera. Excision was stimulated 100-1000 times by induction of single-stranded replication. It was proposed that the presence of single-stranded DNA might increase the chance of alignment of direct repeats. Also deletion via this mechanism only required that the direct repeats are in a single-stranded form. These results also confirm the earlier work performed by Hermann *et al.* (1978) who demonstrated increased deletion formation in fd phage which can replicate via a single-stranded phase. Evidence also exists to suggest a bias for deletion on the lagging strand of the replication fork (Trinh and Sinden, 1991). Replication on the lagging strand is discontinuous in nature (Friedberg, 1992) with regions of single-strand DNA formed between the intiation of Okazaki fragments.

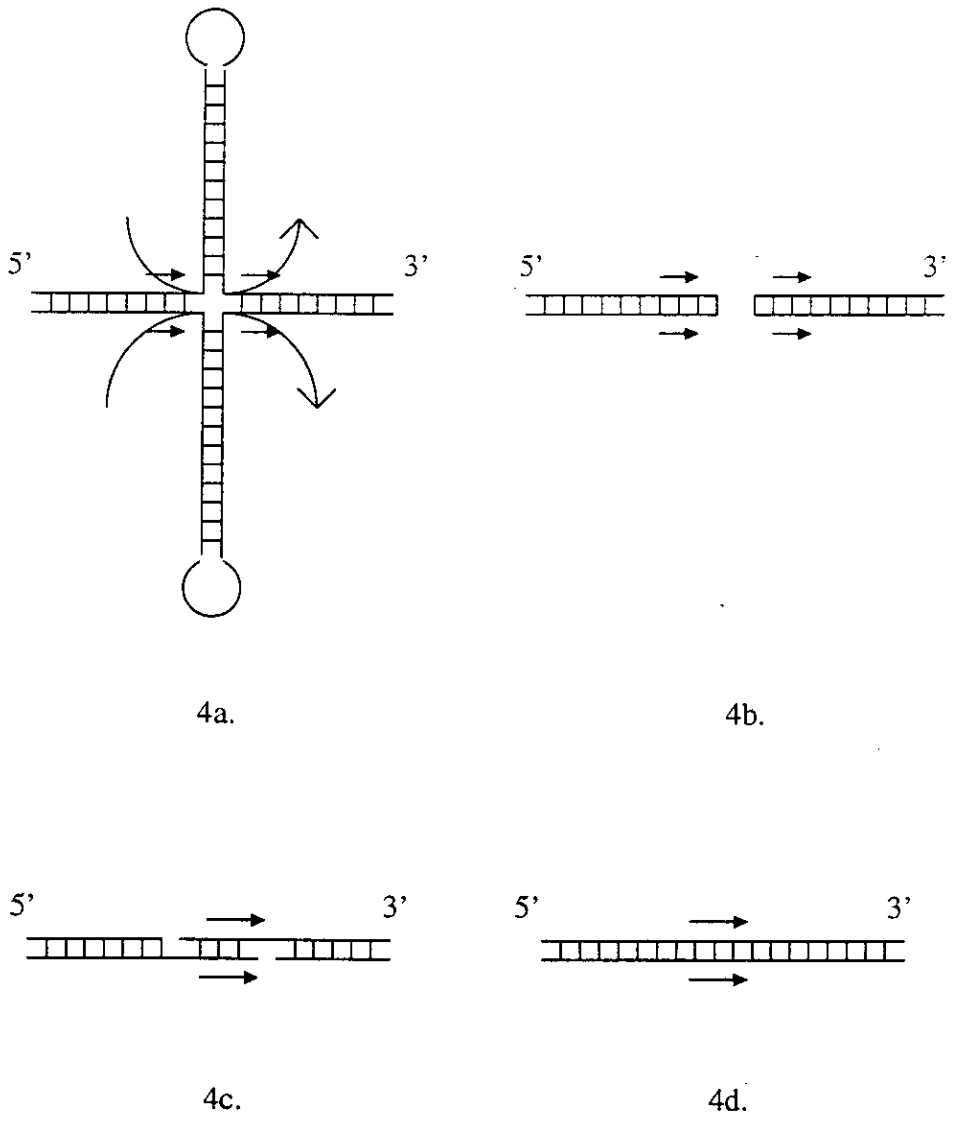
Preference for one of the direct repeats used in palindrome instability to be located just inside the palindrome and one outside indicates that replication is able to

proceed some way into a secondary structure before dissociation of the nascent strand and subsequent annealment at a second copy of the repeat (see earlier section). Also the prevalence of direct repeats as deletion endpoints in many illegitimate events can be explained by the replication slippage model.

There are also the observations that 530 bp and 8400 bp palindromes are only lost from replicating λ phage (Leach and Lindsey, 1986; Shurviton *et al.*, 1987) and transposon excision from a plasmid/phage chimera was highest when replication switched from single-stranded replication to double stranded replication (Brunier *et al.*, 1988). It should be noted however, that not all illegitimate events involve the use of direct repeats or secondary structure (e.g. Benson and Bremer, 1987) so it is difficult to apply the model of replication slippage in these cases.

The second model proposed for illegitimate recombination involves the active cutting and rejoining of DNA strands (fig.4). In the case of palindrome instability this could involve enzymatic cleavage across the base of the palindrome leaving a double strand break. This would be followed by limited 3'-5' exonuclease activity which allows annealing of the direct repeats. Ligation of the two ends would result in one copy of the direct repeats being retained. Many of the enzymes which cut DNA have specific recognition sequences. If these enzymes participated in

Figure 4. Break-join model for deletion formation (after DaGupta *et al.*, 1987).



4a. 4b.

4c. 4d.

4a. Cleavage across the base of the cruciform by an endonuclease.

4b. Flush double-stranded DNA ends result.

4c. Limited 3'-5' exonuclease digestion and annealing of complementary sequences.

4d. Ligation results in one copy of the direct repeats being retained.

illegitimate recombination, then it would be expected that their recognition sequences would frequently occur at or near the sites of deletion formation. Work by Ikeda (1982, 1990) to correlate the site of illegitimate recombination in plasmid/phage fusions with gyrase cleavage is far from conclusive. Glickman and Ripley (1986) suggested that the sequence 5'GATC 3' was located at the preferred sites for deletion formation. GATC is the DNA sequence that is cleaved in methyl-directed mismatch repair (Lu *et al.*, 1984) and this preference may account for the fact that when searching for mutants with altered frequencies of transposon excision Lundblad and Kleckner (1985) isolated mutations in several of the mismatch repair genes.

None of the models described for illegitimate recombination is able to adequately explain all the features of recorded recombination events. It seems likely that given the variety of genome rearrangements which can result from illegitimate recombination and the range of frequencies at which they occur, more than one mechanism may actually operate *in vivo*. What is apparent is that despite many years of research by a number of workers, very little is known about illegitimate recombination as compared to homologous and site-specific recombination.

Aims of thesis.

The aims of this thesis are as follows:-

1. To investigate the deletion of long DNA palindromes from high copy number plasmids and look at the factors affecting deletion.

2. Characterise plasmid derivatives which apparently exhibit increased levels of palindrome stability.

3. Examine the behaviour of high copy number plasmids bearing long palindromes in a variety of *E. coli* backgrounds with a view to the correct choice of strain for the cloning of long palindromes.

4. Look for a phenotype associated with *sbcCD* strains other than those previously documented.

CHAPTER 2
MATERIALS & METHODS

SECTION 1

MATERIALS

Microbiological Strains, Media and Solutions

All bacterial strains, plasmids and sequencing primers used in this work are described in tables 1a, 1b and 1c respectively.

Media

Bacteriological Media

The following quantities are for 1 litre final volumes made up in distilled water and sterilised by autoclaving for 20 minutes at 15 lb in^{-2}.

BBL

10 g trypticase (Baltimore Biological Laboratories), 5 g NaCl, 10 g Bacto-agar (Difco), adjusted to pH 7.2 with NaOH.

BBL Top Agar

As BBL agar, but containing only 6.5 g Bacto-agar (Difco) per litre.

L Agar

10 g Bacto-tryptone (Difco), 5 g yeast extract (Difco), 10 g NaCl, 15 g Bacto-agar (Difco), adjusted to pH 7.2 with NaOH.

L Broth

10 g Bacto-tryptone (Difco), 5 g yeast extract (Difco), 10 g NaCl, adjusted to pH 7.2 with NaOH.

PA Agar

200 ml PA 10x salts, 20 ml PA 100x salts, 2 ml 1% vitamin B1, 20 ml 20% glucose or 10 ml 20% lactose + 10 ml 20% arabinose, 15 g Bacto-agar (Difco).

PA Broth

200 ml PA 10x salts, 20 ml PA 100x salts, 2 ml 1% vitamin B1, 20 ml 20% glucose.

Table 1a. *E. coli* strain genotypes

Strain	Genotype	Reference
JM83	F' *ara* Δ(*lac-proAB*) *rpsL* (Str^R) [φ80 d*lac*Δ(*lacZ*) M15]	1
DL324	JM83+pAC2	D.Leach
DL494	JM83 *sbcC*201	D.Leach
DL733	JM83 Δ*sbcCD*::kan^R	D.Leach
DL887	JM83 *recA*::Cm^R	This work
DL888	DL733 *recA*::Cm^R	This work
DL528	DL494 +pAC2	D.Leach
DL917	JM83+pDLJ2	This work
DL918	DL494+pDLJ2	This work
DL919	DL733+pDLJ2	This work
DL920	DL887+pDLJ2	This work
DL921	DL888+pDLJ2	This work
DL922	JM83 *sfiA*::kan^R	This work
DL923	DL494 *sfiA*::kan^R	This work
DL924	DL922+pDLJ2	This work
DL925	Dl923+pDLJ2	This work
MCS2a2	F*thi*Δ(*lacIPOZYA argF*)U169 *fla relA rpsL araD*139*araB*::+Muc*ts*62/*l* pl(209, U118) DW209*trp-lac*0	2
Mcs2a3	F*thi*Δ(*lacIPOZYA argF*)U169 *fla relA rpsL araD*139*araB*::+Muc*ts*62/*l* pl(209, U118) DW209*trp-lac*0	2
DL943	MCS2a2 Δ*sbcCD*::kan^R	This work
DL944	MCS2a2 *mutS*	This work
DL945	MCS2a2 Δ*sbcCD*::kan^R *mutS*	This work
DL947	MCS2a3 Δ*sbcCD*::kan^R	This work
DL949	JM83+pUC18	This work
DL951	DL733+pAC2	This work
DL952	DL733+pMS7	This work

DL954	DL733+pUC18	This work
DL956	DL887+pUC18	This work
DL958	DL888+pAC2	This work
DL959	DL888+pMS7	This work
DL961	DL888+pUC18	This work
DL473	AB1157 F'*thr*-1*ara*-14*leuB*6Δ(*gpt-proA*)62*lacY*1*tsx*-33*supE*44*galK*2λ⁻rac⁻ *hisG*4*rfbD*1*mgl*-51*rspL*31*kdgK*51*xyl*-5*mtl*-1*argE*3*thi*-1	3
DL591	DL473 *recA*	D.Leach
DL520	DL473 *recA recD*	D.Leach
DL521	DL473 *recA sbcC*	D.Leach
DL522	DL473 *recA recD sbcC*	D.Leach
DL847	K-12 25H28 *recA recD sbcC*	D.Leach

References

1. Yanisch-Perron, C., Viera, J. and Messing, J. (1985) Gene **33**: 103-119
2. Shapiro, J. (1984) Mol. Gen. Genet. **194**: 79-90
3. Howard-Flanders, P., Boyce, R.P. and Theriot, L. (1964) Genetics **49**: 237-246

Table 1b. Plasmids used in this work

Plasmid	Contruct	Reference
pAC2	pUC18+571 bp pal.	1
pMS5	pUC18+571 bp pal.	2
pMS7	pUC18+571 bp pal	2
pUC18	-	3
pDLJ1	pUC118+109 bp pal.	4
pDLJ2	pUC119+109 bp pal.	4

References

1. Chalker, A.F. Ph.D. Thesis (1990) University of Edinburgh.

2. M. Shaw and D. Leach, unpublished.

3. Yanisch-Perron, C., Viera, J. and Messing, J. (1985) Gene **33**: 103-119

Table 1c. Primers used for plasmid DNA sequencing

Nos.	Orientation	Primer Sequence	Position (bp)	Refs.
P2845	Sense	5'-CAATACGCAAACCGCCTC-3'	6-23	This work
596L	Sense	5'-GACTGGAAAGCGGGCA-3'	74-89	-
-40	Anti-sense	5'-GTCGTGACTGGGAAAAC-3'	310-326	1
P38	Sense	5'-TTCCCAACAGTTGCGCAG-3'	413-430	This work
P39	Sense	5'-TCGGGGAAATGTGCGCGG-3'	767-784	This work
F662	Sense	5'-GAGCAACTCGGTCGCCGC-3'	1143-1160	This work
F664	Sense	5'-GGAGCCGGTGAGCGTGG-3'	1584-1600	This work
H303	Anti-sense	5'-GGTGCCTCACTGATTAAG-3'	1719-1736	This work
2975	Sense	5'-TTTCGTTCCACTGAGCGT-3'	1867-1884	This work
F663	Sense	5'-TGGTTTGTTTGCCGGATC-3'	1981-1998	This work
745W	Sense	5'-TATCCGGTAAGCGGCAGG-3'	2352-2369	This work
H304	Anti-sense	5'-GTCCTGTCGGGTTTCGCC-3'	2428-2445	This work

References.

1. -40 primer supplied with Sequenase Version 2.0 sequencing kit from USB.

All other primers designed for this study and synthesized by Oswell, UK.

PA 100X Salts

20 g $MgSO_4$. 7H2O, 0.1 g $FeSO_4$. $7H_2O$, 10 mg $ZnSO_4$. $7H_2O$*, 10 mg $CuSO_4$. $5H_2O$*, 4 mg H_3BO_3*, 4 mg $MnSO_4$. $7H_2O$, 2 mg MoO_3. H_2O*. * Diluted from concentrated stock solutions.

Phage Buffer

3 g KH_2PO_4, 7 g Na_2HPO_4, 5 g NaCl, 1 mM $MgSO_4$, 1 mM $CaCl_2$, 1 ml gelatin (1% w/v).

TM Buffer

10 mM Tris/HCl (pH 7.5), 10 mM $MgSO_4$.

Media Additives

1 M $MgSO_4$ Stock

Made up in distilled water, autoclaved.

20% Arabinose

Made up in distilled water and filter sterilised.

20% Glucose

Made up in distilled water and filter sterilsed.

20% Lactose

Made up in distilled water and filter sterilised.

20% Maltose

Made up in distilled water and filter sterilised.

Ampicillin (100 mg ml^{-1})

Ampiclin (Beecham Pharmaceuticals) was made up in sterile distilled water and stored at -20 °C. It was used at 100 $\mu g\ ml^{-1}$.

Cephalexin (10 mg^{-1})

Cephalexin (gift from Dr.K.Begg) was made up in sterile distilled water and stored at -20 °C. It was used at 10 $\mu g\ ml^{-1}$.

Chloramphenicol (20 mg ml^{-1})

Chloramphenicol (Sigma Chemical Company) was made up in 100% ethanol and stored at -20°C. It was used 50 $\mu g\ ml^{-1}$.

Rifampicin (100 mg ml^{-1})

Rifampicin (Sigma Chemical Company) was made in dimethylformamide and stored in the dark at -20°C. It was used at 150 $\mu g\ ml^{-1}$.

Tetracycline (15 mg ml^{-1})

Tetracycline (Sigma Chemical Company) was made up in 50 % ethanol and stored at - 20°C. It was then used at 10 $\mu g\ ml^{-1}$.

Trimethoprim (5 mg ml^{-1})

Trimethoprim (Sigma Chemical Company) was made up in 100% methanol and stored at 4°C. It was used at 10 $\mu g\ ml^{-1}$.

IPTG (20 mg ml^{-1})

IPTG (isopropyl-thiogalactoside) (Calbiochem) was made up in distilled water and filter sterilised. It was stored at -20°C in 1 ml aliquots.

X Gal (25 mg ml^{-1})

X-gal (5-bromo-4-chloro-3 indolyl-β-D-galactoside) (Calbiochem) was made up in dimethylformamide and stored at -20°C. It was used at 40 $\mu g\ ml^{-1}$.

1% Vitamin B1

Vitamin B1 (Sigma Chemical Company) was made up in distilled water and filter sterilised. It was stored at 4°C.

Solutions

Solutions for Transformaton of *Escherichia coli*

0.1 M $CaCl_2$

Made up in distilled water, autoclaved.

Solutions for UV Mutagenesis of *Escherichia coli*

0.1 M $MgSO_4$

Made up in distilled water, autoclaved.

Solutions for Flow Cytometry

0.1M Tris/HCl (pH 7.4)

0.1 M Tris base, adjusted to pH7.4 with concentrated HCl, autoclaved.

1 M NaCl

Made up in distilled water, autoclaved.

0.1 M $MgCl_2$

Made up in distilled water, autoclaved.

Mithramycin (20 mg ml^{-1})

Mithramycin (Pfizer) was made up in sterile distilled water and stored at -20°C.

Ethidium Bromide (10 mg ml^{-1})

Ethidium bromide (Calbiochem) was made up in sterile distilled water and stored at 4°C.

Staining Solution

0.1M Tris/HCl, 50 µg ml^{-1} mithramycin, 25 µg ml^{-1} ethidium bromide, 25 mM $MgCl_2$ and 100 mM NaCl. Made up fresh and stored at 4°C.

77% Ethanol

Prepared by adding 0.23 volumes of sterile distilled water to 0.77 volumes of ethanol.

Materials for DNA Purification and Manipulation

General Solutions and Materials for DNA Purification

Unless otherwise stated, general laboratory chemicals were purchased from Sigma Chemical Company, Fisons or BDH.

1 M Tris/HCl Stock (pH7.5)

1 M Tris base, adjusted to pH7.5 with concentrated HCl, autoclaved.

0.5 M EDTA (pH8.0)

0.5 M EDTA disodium salt, adjusted to pH8.0 with glacial acetic acid, autoclaved.

TE Buffer Stock (10X)

100 mM Tris, 10 mM EDTA, adjusted to pH7.5 with concentrated HCl, autoclaved.

Solutions for Phenol-Chloroform and Chloroform Extraction

Phenol-Chloroform

Distilled, liquified 88% phenol (Rathburn Chemicals) was stored in 25 ml aliquots at -20°C, and protected from the light in 25 ml polypropylene tubes. After thawing at room temperature 0.1% (w/v) 8-hydoxyquinoline (Sigma Chemical Company) was added and the phenol equilibrated using TE buffer. Initially this was performed by emulsifying the phenol with 10X TE buffer for 5 mins, followed by centrifugation at 4.5 krpm in an MSE Centaur-2 bench centrifuge. The aqueous layer was removed, and equilibration repeated another two times using 1x TE buffer. The aqueous layer was then removed and the phenolic layer combined with an equal volume of 24:1 chloroform-isoamylalcohol and centrifuged at 4.5 krpm in an MSE Centaur-2 bench centrifuge. The aqueous layer was removed and replaced with 2 ml 1X TE buffer

0.2% (v/v) β-mercaptoethanol (Sigma Chemical Company). The phenol-chloroform was protected from the light, stored at 4°C and used within 1 month.

24:1 Chloroform-Isoamylalcohol

24 volumes of chloroform were combined with 1 volume of isoamylalcohol and stored away from light.

Solutions for Ethanol and Isopropanol Precipitation

3 M Sodium Acetate (pH5.3)

Sodium acetate solution (3 M in acetate) was prepared by adding 0.19 volumes of sterile 3 M acetic acid to 0.81 volumes of sterile 3 M sodium acetate, then autoclaved.

3 M Sodium Acetate (pH7.0)

3 M sodium acetate, adjusted to pH7.0 with glacial acetic acid and then autoclaved.

70% Ethanol

Prepared by adding 0.3 volumes of sterile distilled water to 0.7 volumes of ethanol.

Preparation of Dialysis Tubing

Dialysis tubing was cut into lenghts of 20 cm and boiled for 10 mins in 2 litres of 2% sodium bicarbonate and 1 mM EDTA. It was then rinsed thoroughly in distilled water. After cooling, the tubing was stored at 4°C submerged in 1 mM EDTA 50% (v/v) ethanol. Before use the tubing was washed inside and out with sterile distilled water.

Materials for Purification of DNA from Agarose Gels and Solutions

Purification of DNA from agarose gels and solutions was performed using GENECLEAN and MERmaid kits (BIO101). The solutions used were those provided by the manufacturer and are listed for reference.

6 M NaI Stock

6 M NaI, protected from the light and stored at 4°C.

GLASSMILK

Suspension of silica matrix in distilled water, stored at 4°C.

NEW Wash

20 mM Tris/HCl (pH7.2), 0.2 M NaCl, 2 mM EDTA, 50% ethanol, stored at -20°C.

High Salt Binding Soluton.

Saturated solution of sodium perchlorate, stored at room temperature.

GLASSFOG

Suspension of fine silica-based matrix in sterile distilled water, stored at room temperature.

Ethanol Wash

20 mM Tris/HCl (pH7.2), 0.2 M NaCl, 2 mM EDTA, 90% ethanol, stored at room temperature.

Solutions for Plasmid DNA Purification

Small Scale Method

TEG

0.9% glucose, 25 mM Tris/HCl (pH7.5), 10 mM EDTA, filter sterilised and stored at 4°C.

Alkaline SDS

0.2 M NaOH, 1% SDS (sodium dodecyl sulphate), freshly prepared.

Potassium Acetate

Potassium acetate (3 M in potassium, 5 M in acetate) was prepared by adding 0.4 volumes of sterile 5 M acetic acid to 0.6 volumes of sterile 5 M potassium acetate. autoclaved and stored at 4°C.

TE RNase A

10 mM Tris/HCl (pH7.5), 1 mM EDTA, 20 µg ml-1 RNase A. Prepared by adding 100 µl RNase A stock (10mg ml^{-1}) to 50 ml sterile 1X TE buffer. Stored at 4°C.

Large Scale Method

Large scale plasmid DNA preparation was performed using a Qiagen plasmid midi kit. The solutions were those provided by the manufacturer and are listed for reference.

Buffer P1

100 µg ml^{-1} RNase A, 50 mM Tris/HCl, 10 mMEDTA stored at 4°C.

Buffer P2

200 mM NaOH, 1% SDS, stored at room temperature.

Buffer P3

3 M potassium acetate (pH5.5) stored at 4°C.

Buffer QBT

750 mM NaCl, 50 mM MOPS (pH7.0), 15% ethanol, 0.15% Triton X-100, stored at room temperature.

Buffer QC

1 M NaCl, 50 mM MOPS (pH7.0), 15% ethanol, stored at room temperature.

Buffer QF

1.25 M NaCl, 50 mM Tris/HCl (pH8.5), 15% ethanol, stored at room temperature.

Caesium Chloride Purified DNA

1X STE

100 mM NaCl, 10 mM Tris/HCl (pH8.0), 1 mM EDTA, autoclaved and stored at 4°C.

Lysozyme

Lysozyme (Sigma Chemical Company) dissolved in 10 mM Tris/HCl (pH8.0), made fresh.

Enzymes and Buffers for DNA Manipulation

Restriction Endonuleases

All restriction endonucleases used in this work are described in table 2.

DNA Sequencing - Chain Termination Method

DNA sequencing was performed using a Sequenase v.2.0 sequencing kit (United States Biochemical). With the exception of [α^{35}S]dATP (Amersham) and the sequencing primer, the solutions used were those provided by the manufacturer and are listed for reference. All solutions were stored at -20°C.

Sequenase Buffer (5X)

200 mM Tris/HCl (pH7.5), 100 mM $MgCl_2$, 250 mM NaCl.

Enzyme Dilution Buffer

10 mM Tris/HCl (pH7.5), 5 mM DTT, 0.5 mg ml^{-1} BSA.

DTT

0.1 M DTT (dithiothreitol), prepared in distilled water.

Labelling Mix (5X)

7.5 µM dGTP, 7.5 µM dCTP, 7.5 µM dTTP.

Table 2. Restriction enzymes used in this work

Enzyme	Target Site	Supplier	Buffer
*Afl*III	AC/PuPyGT	New England Biolabs	3
*Alu*I	AG/CT	New England Biolabs	1
*Alw*NI	CAGNNN/CTG	New England Biolabs	4
*Eco*RI	G/AATTC	Northumbria Biologicals	6
*Hae*III	GG/CC	New England Biolabs	2
*Hha*I	GCG/C	New England Biolabs	4
*Hin*dIII	A/AGCTT	Northumbria Biologicals	10
*Sac*I	GAGCT/C	Northumbria Biologicals	2
*Sca*I	AGT/ACT	Northumbria Biologicals	6

Buffer composition

NEB 1. 10 mM Bis Tris Propane-HCl, 10 mM Mg Cl_2, 1 mM dithiothreitol.

NEB 2. 10 mM Bis Tris Propane-HCl, 10 mM Mg Cl_2, 1 mM dithiothreitol.

NEB 3. 100 mM Na Cl, 50 mM Tris-HCl, 10 mM Mg Cl_2, 1 mM dithiothreitol.

NEB 4. 50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM dithiothreitol.

NBL 2. 33 mM Tris-acetate (pH 8.2), 66 mM potassium acetate, 10 mM magnesium acetate, 0.5 mM dithiothreitol.

NBL 6. 50 mM Tris-HCl (pH 7.8), 100 mM NaCl, 10 mM $MgCl_2$, 1 mM dithiothreitol.

NBL 10. 50 mM Tris-HCl (pH 8.3), 50 mM NaCl, 10 mM $MgCl_2$, 1 mM dithiothreitol.

ddGTP Termination Mix

80 µM dGTP, 80 µM dATP, 80 µM dCTP, 80 µM dTTP, 8 µM ddGTP, 50 mM NaCl.

ddATP Termination Mix

80 µM dGTP, 80 µM dATP, 80 µM dCTP, 80 µM dTTP, 8 µM ddATP, 50 mM NaCl.

ddCTP Termination Mix

80 µM dGTP, 80 µM dATP, 80 µM dCTP, 80 µM dTTP, 8 µM ddCTP, 50 mM NaCl.

ddTTP Termination Mix

80 µM dGTP, 80 µM dATP, 80 µM dCTP, 80 µM dATP, 8 µM ddTTP, 50 mM NaCl.

Stop Solution

95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol.

Maxam-Gilbert Chemical Sequencing

10 % Formic Acid

Prepared by adding 0.1 volumes of formic acid to 0.9 volumes of sterile distilled water.

1 M Piperidine

Made up in sterile distilled water just prior to use.

30 % NaOH

Made up in sterile distilled water

1 % SDS

Made up in sterile distilled water.

$KMnO_4$ (2 mg ml^{-1})

Made up in sterile distilled water and stored in the dark for up to one week at 4°C. Used at 20 µg ml^{-1}.

Other Eznymes

Incubation buffers were those supplied by the manufacturer.

Bacteriophage T4 DNA Ligase

Bacteriophage T4 DNA ligase (New England Biolabs) was incubated in 50 mM Tris/HCl (pH7.5), 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 25 µg ml^{-1} BSA.

Calf Intestinal Alkaline Phosphatase (CIP)

Calf Intestinal Alkaline Phosphatase (Boehringer Mannheim) was incubated in 1 mM $ZnCl_2$, 1 mM $MgCl_2$, 10 mMTris/HCl (pH8.3).

Klenow Enzyme

Klenow Enzyme (Boehringer Mannheim) was incubated in 10 mM Tris/HCl (pH7.5), 5 mM $MgCl_2$ and 7.5 mM DTT.

Other Solutions

dNTP Stocks

dNTPs (Sigma Chemical Company) were prepared in sterile distilled water at a concentration of 50 mM and 2 mM. Stored at -20°C.

BSA (20 mg ml^{-1})

Bovine Serum Albumin (Boehringer Mannheim) was stored at -20°C.

Solutions for Gel Electrophoresis

Agarose Gel Electrophoresis

TAE Gel Buffer Stock

0.8 M Tris acetate, 20 mM EDTA (pH8.0).

TAE Gel-Loading Sample Buffer Stock (5X)

0.2 M Tris acetate, 0.25 M EDTA (pH8.0), 0.2% bromophenol blue, 15% Ficoll 400.

Ethidium Bromide (10 mg ml^{-1})

Ethidium bromide (Sigma Chemical Company) was prepared in sterile 1X TE buffer, stored protected from the light at 4°C. Used at 0.5 μg ml^{-1}.

TBE Gel Buffer Stock (10X)

0.89 M Tris borate, 10 mM EDTA (pH8.0)

Chloroquine (10 mg ml^{-1})

Chloroquine (Sigma Chemical Company) prepared in sterile distilled water, stored away from light at 4°C. Used at a concentration of 1.8 μg ml^{-1} or 20 μg ml^{-1}.

Polyacrylamide Gel Electrophoresis

TBE Gel Buffer stock

0.89 m Tris borate, 20 mM EDTA (pH8.0).

Formamide-EDTA Gel-Loading Sample Buffer

98% formamide, 0.1% bromophenol blue, 0.1% xylene cyanol FF, 10 mM EDTA (pH8.0).

40% Acrylamide Stock Solution

38% (w/v) acrylamide, 2% (w/v) N-N-methylene bis-acrylamide (Northumbria Biologicals), protected from light at 4°C.

0.5X TBE 12% Acrylamide Solution

46 g urea, 5 ml 10X TBE buffer, 30 ml 40% acrylamide stock, distilled water to 100 ml total. Degassed and stored protected from the light at 4°C.

Long Ranger Gel Solution

Long Ranger (AT Biochem) 50% solution, stored at room temperature.

1X TBE 6% Long Ranger Gel Solution for Sequencing Gels

21 g urea, 6 ml Long Ranger, 5 ml 10X TBE buffer, distilled water up to 50 ml total volume. Solution prepared just before use.

TEMED

TEMED (N-N-N'-N'-tetra-methyl-1, 2-diamino-ethane)(Sigma Chemical Company). Stored protected from the light at 4°C.

10% AMPS

10% AMPS (ammonium persulphate) was freshly prepared in sterile distilled water.

Gel Fix

10% (v/v) glacial acetic acid, 10% (v/v) methanol in distilled water.

SECTION 2

METHODS

Microbiological Methods

Bacterial Methods

Storage of Bacteria

For short term storage, bacteria were kept at 4°C on L-agar plates sealed with parafilm. Permanent stocks were prepared by adding 5 drops of sterile 100% glycerol to 1 ml of a stationary phase culture in an Eppendorf tube. This was sealed with parafilm and stored at -70°C.

Growth of Bacteria

Temporary stocks were generated by streaking permanent stocks to single colonies on L-agar plates, which were incubated at 37°C overnight unless otherwise stated. Overnight cultures were grown by innoculating a single colony into L-broth and shaking at 37°C unless otherwise stated.

Test for UV Sensitivity

This method was used to test the sensitivity of *E. coli recA* strains to ultra-violet (u.v.) light (see Chapter 3). An overnight culture was diluted 40-fold in to L-broth and grown shaking at 37°C to OD_{650}=0.5. The culture was then serially diluted in ice-cold L-broth and 100 µl of the 10^{-3} and 10^{-4} dilutions were spread on two sets of fresh, dry L-agar plates. One set of plates was incubatedat 37°C overnight. The

other set was exposed for 20 seconds to an ultra-violet lamp (Ultraviolet Products Incorporated) giving 1 J m^{-2} s^{-1} illumination, wrapped immediately in foil (to prevent DNA photolyase activation) and incubated at 37°C overnight. Colonies were counted and the strain showing about 2% survival was chosen (McGraw and Marinus. 1980).

AraB-LacZ Fusion Assay

Plate Assay

A number of independent cultures of *araB-lacZ* fusion strains were grown overnight at 30°C in PA medium with 0.4% glucose. Two dilutions of overnight cells were plated on fresh plates containing 0.2% lactose and 0.2% arabinose. The plates were incubated at 30°C and the number of fusion colonies appearing every day were scored for ~30 days (Shapiro, 1984).

Liquid Assay

A number of cultures were grown overnight in PA medium with 0.4% glucose. The cultures were then split in two, one half being continuously incubated with aeration, the other half unaerated. Each day 100 µl aliquots of the cultures were plated on PA agar with 0.2% lactose and 0.2% arabinose to assay for fusion colonies which appear after two days incubation on the plates. The liquid cultures were also assayed for viable count by plating dilutions on PA agar with 0.4% glucose (Maenhaut-Michel and Shapiro, 1994).

P1 Transduction

The strain to be transduced was grown overnight in L-broth supplemented with 2.5 mM $CaCl_2$. 1 ml aliquots of the culture were centrifuged at 15 krmp to pellet the cells (Sorvall Microspin 24 centrifuge). The pellet was then resuspended in 0.1 ml of various dilutions of the P1 lysate to be used. The cells were then incubated at 37°C for 20 minutes. 0.8 ml of L-broth supplemented with 2mM sodium citrate were then added and incubation continued at 37°C for a further 30-60 minutes. 0.1 ml aliquots were then plated out on the appropriate selective medium and incubated at 37°C overnight (unless otherwise stated).

Potential transductants were streaked out on selective medium and any additional tests for the loss of the gene of interest carried out.

Flow Cytometry

Method previously described by Skarstad and Boye (1993).

Overnight cultures of the strains to be studied were diluted 1:1000 into minimal medium supplemented with 0.5% casamino acids and 0.2% glucose. The cultures were incubated at 37°C until an OD_{750}=0.15 was reached. 1.5 ml of the culture was removed and cooled on ice in an Eppendorf. To the remaining culture 10 μg ml^{-1} cephalexin and 150 μg ml^{-1} rifampicin were added and incubation at 37°C continued for 4-6 hours. 1.5 ml aliquots of the cultures were centrifuged for 1 minute at 15 krpm (Sorvall Microspin 24 centrifuge) along with the samples stored on ice. The pellets were then resuspended in 100 μl 1x TE and 1 ml 77% ethanol added. The

samples were then stored at 4°C. Prior to staining for flow cytometry the cells were pelleted at 15 krpm (Sorvall Microspin 24 centrifuge) for 1 minute and then washed twice in 0.1M Tris/HCl (pH7.4). The pellet was then resuspended in 1ml of staining solution. Flow cytometric measurements were performed with an Argus flow cytometry (Skatron, Tranby, Norway). In this instrument, cells pass one at a time through a beam of excitation light. Each cell gives rise to a pulse of fluorescence light, the size of the pulse being proportional to the amount of mithramycin bound to the DNA. The excitation wavelength used did not excite the ethidium bromide. The fluorescence from and the light scattered by each bacterial cell was proportional to the mass of the cell. The pulses of light are sized and stored by a mulitchannel pulse height analyzer. The multichannel analyzer thus accumulates a DNA histogram (a distribution of bacteria with regard to their DNA content). Cells are measured at a rate of up to 10^4 cells sec^{-1}.

Plasmid Methods

Maintenance of Plasmids

Plasmids were maintained by supplementing media with antibiotics.

Transformation

This method is derived from Mandel and Higa (1970).

Preparation of Competant Cells

A fresh overnight culture was diluted 1:100 in L-broth and grown shaking at 37°C to an OD_{650}=0.35-0.45. The culture was chilled on ice for 15 minutes and 20 ml were centrifuged at 4.5 krpm for 5 minutes at 4°C (MSE Centaur-2 bench centrifuge). The cell pellet was resuspended in 2 ml ice-cold 0.1 M $CaCl_2$ and incubated on ice for 20 minutes. The sample was again centrifuged at 4.5 krpm for 5 minutes at 4°C and resuspended gently in 400 µl of ice-cold $CaCl_2$. The competant cells were used immediately.

Transformation

A 200 µl aliquot of competant cells was added to 50-100 ng of plasmid DNA (suspended in 1-10 µl of 1X TE buffer or sterile distilled water) and incubated on ice for 30 minutes. The cells were then heat-shocked fo 2 minutes at 42°C in a waterbath then rapidly cooled on ice for 2 minutes. 10 µl and 100 µl aliquots of the transformed cells were then spread on plates containing the appropriate antibiotic and incubated at 37°C overnight unless otherwise stated.

Detection of Inserts in pUC Based Plasmids

L-agar was supplemented with X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside) at a final concentration of 40µg ml^{-1} and ampicillin at a final concentration of 100 µg ml^{-1}. 40 µl of IPTG (isopropyl-thiogalactaside) at a concentration of 20 mg ml^{-1} were spread on the plates. Blue colour represents those cells bearing plasmids which have

deleted the palindrome, white cells, those which still carry the palindrome (see Chapter 3).

DNA Purification and Manipulation

General Methods of DNA Purification

Phenol-Chloroform and Chloroform Extraction

Successive phenol-chloroform and chloroform extractions were generally sufficient to purify DNA from contaminating protein.

Phenol-Chloroform

Phenol-chloroform extraction was carried out in Eppendorf tubes. An equal volume of equilibrated phenol-chloroform was added to the DNA solution and the two were mixed by vortexing briefly. The two phases were separated by centrifugation at 15 krpm for 15 minutes (Sorvall Microspin 24 centrifuge). The upper (aqueous) layer was then transferred to a fresh Eppendorf tube.

Chloroform-Isoamylalchohol Extraction

Chloroform extraction was peformed in the same way as phenol-chloroform extraction, except that 24:1 chloroform-isoamylalchohol was used in the place of phenol-chloroform, and the phases were separated by only 1 minute centrifugation at 15 krpm.

Ethanol and Isopropanol Precipitation

Ethanol precipitation was normally used to concentrate and purify DNA. Isopropanol precipitation was used where a large volume was impractical or where it was desirable to avoid the co-precipitation of RNA.

Ethanol Precipitaton

Ethanol precipitaton was caried out in 15 ml corex tubes or Eppendorf tubes. 0.1 volumes of 3 M sodium acetate pH5.3 were added to 0.9 volumes of DNA solution, followed by two volumes of freezer-cold (-20°C) ethanol. Where the concentration of EDTA in the DNA solution exceeded 10-15 mM, 3 M sodium acetate pH7.0 was used (high concentrations of EDTA precipitate from solution at acid pH). The contents of the tubes were mixed by inversion and incubated at -20°C for at least one hour to allow precipiatation. The DNA was then pelleted by centrifugation at 15 krpm for 20 minutes at 4°C (Sorvall RC5B centrifuge with SS34 rotor or Sorvall Microspin 24 centrifuge). The ethanol was discarded, the tube filled with freezer-cold 70% ethanol and centrifuged again at 15 krpm for 10 minutes at 4°C. The 70 % ethanol was discarded, the pellet air dried for 15 minutes and redissolved in 1X TE buffer or sterile distilled water.

Isopropanol Precipitation

Isopropanol precipitation was carried out in a similar way to ethanol precipitation, except that 0.7 volumes of isopropanol at room temperature were used in place of 2 volumes of freezer-cold ethanol. The tube was incubated at room temperature for 15

minutes to allow precipitation, and centrifugation (15 krpm for 20 minutes, as above) was also carried out at room temperature. The pellet was rinsed with freezer-cold 70% ethanol and recentrifuged as described above.

Purification of DNA From Agarose Gels and Solutions

Purification of DNA from agarose gels and solutions was performed using GENECLEAN and MERmaid kits (BIO101). The procedures are based on the method of Vogelstein and Gillespie (1979).

GENECLEAN Proceedure

The band of interest was excised from the agarose gel and its weight determined in a tared Eppendorf tube. 3 volumes (v/w) of 6 M NaI stock were added and the agarose dissolved by 5 minute incubation in a 50°C waterbath, with occasional mixing by tube inversion. Where the DNA was in aqueous solution, 3 volumes (v/v) of 6 M NaI stock were added without incubation at 50°C. 5 µl of GLASSMILK suspension were added, mixed by vortexing and incubated on ice for 5 minutes. The silica matrix with bound DNA was then pelleted by centrifugation for 30 seconds at 15 krpm (Sorvall Microspin 24 centrifuge) and the supernatant discarded. The pellet was resuspended in 500 µl of freezer-cold NEW wash and centrifuged for 30 seconds at 15 krpm (Sorvall Microspin 24 centrifuge). The supernatant was discarded and the pellet washed a further two times with 500 µl NEW wash. After the supernatant was removed the pellet was air dried for 10 minutes, resuspended in 5 µl of sterile distilled water and incubated in a 50°C waterbath for 3 minutes. It was then

centrifuged for 1 minute at 15 krpm (Sorvall Microspin 24 centrifuge) and the supernatant containing the eluted DNA was transferred to a fresh Eppendorf tube. A second elution step was performed and the two supernatants combined.

MERmaid Procedure

This method was used to purify DNA fragments of less than 200 bp from agarose gels or solutions. The band of interest was excised from the agarose gel and its weight determined in a tared Eppendorf tube. 3 volumes (v/w) of high salt binding solution were added wth 8 µl of GLASSFOG suspension. The agarose gel was dissolved by continuous vortexing for 15 minutes. If the DNA was in solution, 3 volumes (v/v) of high salt binding solution and 8 µl of GLASSFOG were added, then the tube was incubated at room temperature for 5-15 minutes. The fine silica matrix with DNA bound was the pelleted by centrifugation for 30 seconds at 15 krpm (Sorvall Microspin 24 centrifuge) and the supernatant dicarded. The pellet was the resuspened in 300 µl ethanol wash by vortexing and centrifuged fo 15 seconds at 15 krpm. The supernatant was discarded and the pellet washed two more times with 300 µl ethanol wash. After the third wash the pellet was air dried for 10 minutes, resuspended oin 5-10 µl of sterile distilled water and incubated at room temperature for 5 minutes. It was then centrifuged for 2 minutes at 15 krpm (Sorvall Microspin 24 centrifuge) and the supernatant containing the eluted DNA transferred to a fresh Eppendorf tube. A second elution step was performed and the two supernatants combined.

Plasmid DNA Purification

Small Scale

This method was derived from Birrnboim and Doly (1979). 1.5 ml of a fresh overnight culture were transferred to an Eppendorf tube and centrifuged fo 30 seconds at 15 krpm (Sorvall Microspin 24 Centrifuge). The supernatant was discarded, the pellet centrifuged again to for 5 seconds and the remaining liquid removed. The pellet was resuspended in 100 µl ice-cold TEG by vortexing and the tube incubated at room temperature for 5 minutes. 200 µl of freshly prepared alkaline SDS were added, the tube inverted five times to mix the solutions and the tube incubated on ice for 5 minutes. 150 µl of ice-cold potassium acetate were added and the tube vortexed upright for 5 seconds. After a 5 minute incubation on ice, the cell debris was pelleted by centrifugation at 15 krpm for 5 minutes at 4°C (Sorvall Microspin 24 centrifuge). 400 µl of supernatant containing the plasmid DNA were transferred to a fresh Eppendorf tube and extracted with equal volumes of phenol-chloroform and chloroform-isoamylalchohol. Two volumes of 100% ethanol (at room temperature) were added, the tube was inverted to mix and incubated at room temperature for 5-10 minute. The DNA was pelleted by centrifugation at 15 krpm for 10 minutes at 4°C (Sorvall Microspin 24 centrifuge). The supernatant was discarded and the pellet was rinsed with 1 ml freezer-cold 70% ethanol. After further centrifugation at 15 krpm for 5 minutes at 4°C, the 70% ethanol was removed and the

pellet was air dried for 15 minutes. The plasmid DNA was redissolved in 50µl TE-RNaseA and stored at -20°C.

Large Scale Method

The large scale purificaton of plasmid DNA was performed using a QIAGEN plasmid midi kit. A 40 ml fresh overnight culture was centrifuged at 10 krpm for 10 minutes at 4°C (Sorvall RC5B centrifuge with SS34 rotor). The cell pellet was resuspended in 4 ml buffer P1 and transferred to a 30 ml glass corex tube. 4 ml buffer P2 were added, mixed gently by inversion and incubated at room temperature for 5 minutes. 4 ml of ice-cold buffer P3 were added, the solution mixed by inverting five times and the tube incubted on ice for 15 minutes. The sample was then mixed by inverting once more and then centrifuged at 16 krpm for 30 minutes at 4°C (Sorvall RC5B centrifuge with SS34 rotor). The supernatant was promptly removed and applied to a QIAGEN -tip 100 column which had previously been equilibrated with 4 ml buffer QBT. After the column had been emptied by gravity flow, it was then washed twice with 10 ml buffer QC to remove contaminants in the DNA preparation from the QIAGEN column. After the wash buffer had completely emptied, the DNA was eluted from the column with 5 ml buffer QF and allowed to empty in to a 15 ml glass corex tube. 3.5 ml (0.7 volumes) isopropanol (at room temperature) were added to the DNA solution and the two mixed by inversion. The tube was then centrifuged at 12 krpm for 30 minutes at 4°C (Sorvall RC5B centifuge with RC5B rotor). The supernatant was discarded and the DNA pellet was washed in freezer-cold 70% ethanol. After further centrifugation at 12 krpm for 10 minutes at 4°C, the

70% ethanol was removed and the DNA pellet was air dried for 15 minutes. The plasmid DNA was redissolved in 100 µl 1X TE buffer.

Caesium Chloride Purified DNA

A 500 ml fresh overnight culture was centrifuged at 8 krpm for 15 minutes at 4°C (Sorvall RC5B centrifuge with GSA rotor). The supernatant was discarded and the pellet resuspended in 250 ml ice-cold 1X STE. The cell suspension was centrifuged at 8 krpm for 15 minutes at 4°C to pellet the cells. The supernatant was discarded and the plasmid DNA extracted with a scaled-up version of the small scale plasmid method described earlier. The pellet was resuspended in 18 ml TEG followed by the addition of 2 ml lysozyme (10 mg ml^{-1} in 10 mM Tris/HCl pH8.0) and 40 ml freshly prepared alkaline SDS. The suspension was mixed by inverting and incubated at room temperatre for 5-10 minutes. 20 ml ice-cold potassium acetate were added, the sample mixed by shaking to remove the two phases and then incubated on ice for 10 minutes. The cell debris was pelleted by centrifugation at 4 krpm for 15 minutes at 4°C (Sorvall RC5B centrifuge with GSA rotor) and the supernatant transferred to a fresh centrifuge bottle. 0.6 volumes of isopropanol (at room temperature) were added, the sample mixed by inversion and incubated at room temperature for 10 minutes. The sample was the centrifuged at 6 krpm for 15 minutes at room temperature (Sorvall RC5B centrifuge with GSA rotor). The supernatant was discarded and the DNA pellet washed with 50 ml 70% ethanol (at -20°C) and centrifuged at 6 krpm for 10 minutes at 4°C. The ethanol was discarded, the pellet air dried for 10 minutes and the redissolved in 9 ml 1X TE buffer. 9 g CsCl

(Boehringer Mannheim) were added and the DNA solution warmed to 37°C until the CsCl had dissolved. 0.9 ml ethidium bromide (10 mg ml^{-1}) were added and the solution transferrred to a centrifuge tube which was sealed, crimped and centrifuged at 45 krpm fot 48 hours at 18°C (Sorvall ultracentrifuge with a Ti50 rotor). The plasmid DNA band was viualised using an ultraviolet lamp and removed from the centrifuge tube using a hypodermic needle and syringe. The volume of sample removed was made up to a total of 9 ml with 1X TE, 9 g of CsCl and 0.9 ml of ethidium bromide added as before. The sample was centrifuged again for 48 hours and the plasmid DNA isolated as previously described. The DNA solution was transferred to a screw cap tube and an equal volume of ice cold buffered butanol added. The tube was shaken vigorously for several minutes then stored on ice until the two liquid phases has separated out. The butanol layer was dicarded and an equal volume of fresh ice-cold butanol added to the plasmid fraction. The process was repeated approximately ten times until the ethidium bromide was removed from the solution. The plasmid fraction from the final extraction was placed in dialysis tubing and the dialysed in 2 litres 1x TE buffer at 4°C overnight, with the buffer being changed once. After dialysis the plasmid solution was transferred to an Eppendorf tube and 0.1 volumes 3 M sodium acetate and 2 volumes of freezer-cold ethanol added. The DNA was precipitated at -20°C for at least one hour then the DNA pelleted by centrifugation at 4°C, 15 krpm for 20 minutes (Sorvall Microspin 24 centrifuge). The ethanol was discarded, the tube filled with 70% ethanol and then centrifuged at 4°C, 15 krpm for 10 minutes (Sorvall Microspin 24 centrifuge). The

ethanol was dicarded, the pellet air dried and then resuspended in 100 µl 1x TE buffer.

DNA Manipulation

Restriction Digests

The restriction endonucleases used in this work are shown in table 2. Restriction digests were performed in 10-40 µl of the appropriate incubation buffer. The buffers were supplied by the manufacturer as 10X stock solutions and 0.1 volumes were added to the DNA solution to prepare the incubation buffer. This was supplemented by BSA (bovine serum albumin) to a final concentration of 100 µg ml^{-1}. A five fold excess of restriction enzyme (5u μg^{-1} DNA) was added; this had been experimentally determined to be the optimum ratio of enzyme to DNA. The sample was then incubated for two hours at 37°C unless otherwise indicated.

Radiolabelling of DNA

Radiolabelling of DNA was carried out in a total volume of 30 µl. 22 µl of the appropriate restriction digest was mixed with 3 µl of Klenow buffer. 1µl (10 µCi) [α^{32}P] dATP was added with 1µl of 2 mM stock solutions of the appropriate non-radioactive dNTP's, as determined by the restriction endonuclease used to cleave the DNA. 1 µl Klenow enzyme (1unit) was added and the reaction was incubated at room temperature for 15 minutes. A further 1 µl Klenow enzyme was added and the

incubation continued for a further 10 minutes at room temperature. 1 µl of 50 mM stock solutions of dGTP, dATP, dCTP and dTTP were added as a chase (to ensure flush terminii were produced) and the incubation continued at room temperature for a further 5 minutes. The DNA was purified using the GENECLEAN prior to use in Maxam-Gilbert sequencing.

DNA Sequencing

Chain Termination Method

DNA sequencing was performed using a Sequenase v2.0 DNA sequencing kit (United States Biochemical). The method is derived from Sanger *et al.* (1977) and uses dideoxy chain-terminating reactions. 3-5 µg of plasmid DNA were purified from aqueous solution using the GENECLEAN proceedure and resuspended in 9 µl of distilled water. This proceedure was performed for all DNA samples except those purified by the caesium chloride proceedure. 1 µl (~0.5 µg) sequencing primer (Oswell DNA service) was added to the DNA in an Eppendorf the. The DNA was then denatured by incubation in a waterbath at 100°C for 4 minutes followed by rapid cooling on dry ice. The labelling reaction shown below was prepared in an Eppendorf tube:

1. DTT	2 µl
2. [α^{35}S] dATP (Amersham)	0.5 µl (5 µCi)
3. Diluted labelling mix (1x in distilled water)	0.7 µl

4. Sequenase buffer (5x) 2 µl

5. Diluted Sequenase enzyme (1/8 in ice-cold enzyme dilution buffer) 2.5 µl

The labelling reaction was kept on ice while the template-primer mix was thawed and immediately centrifuged at 15 krpm for 10 seconds to remove the DNA solution from the walls of the tube. The template-primer mix was then added to the labelling reaction and incubated at 20°C for 4 minutes. 4 µl aliquots of the labelling mix were then transferred to four Eppendorf tubes, each containing 2 µl of either ddG, ddA, ddT or ddC termination mix; the tubes had been pre-warmed at 37°C (or 42°C) for 1 minute before the addition of 4 µl of labelling reaction. The termination reactions were then incubated at 37°C for 3 minutes. Where the template was capable of forming secondary structures (and thereby causing band compressions) the termination reactions were performed at 42-45°C. The reactions were halted by the additions of 4 µl of stop solution to each tube and placing the sample on ice. The sequencing reactions were stored at -20°C until electrophoresis.

When sequencing PCR products the Sequenase v2.0 PCR product sequencing kit was used. 5 µl of PCR product were incubated at 37°C with 1 µl Exonuclease I and 1 µl Shrimp Alkaline Phosphatase for 30 minutes (supplied with kit). The reaction was then incubated at 80°C for 30 minutes to heat inactivate the enzymes. The remainder of the protocol is the same as that for normal sequencing except that the Sequenase

enzyme is pre-diluted in a high glycerol solution and this results in Tris-Taurine-EDTA buffer (supplied with the kit) having to be used in the sequencing gel.

Chemical Cleavage Sequencing

This method is derived from that previously described by Maxam amd Gilbert (1977).

A+G Reaction

10 µl of radio-labelled DNA were mixed with 4 µl of sonicated calf thymus DNA. 3 µl of 10% formic acid were then added and the reaction incubated at 37°C for 7 minutes. The sample was then cooled on ice and 150 µl of ice cold 1 M piperidine added, followed by incubation at 90°C for 30 minutes. The sample was then cooled to room temperature, 1.2 ml of butan-1-ol added and the reaction vortexed for 30 seconds to mix.

A>C Reaction

5 µl of radio-labelled DNA were mixed with 4 µl of sonicated calf thymus DNA. 1 µl 30% NaOH was added and the reaction incubated at 90°C for 6 minutes. The reaction was then cooled on ice and 150 µl of ice cold 1 M piperidine added. The reaction was then incubated at 90°C for 30 minutes followed by cooling to room temperature. 150 µl 70% ethanol and 1.2 ml butan-1-ol were added and the sample vortexed for 30 seconds to mix.

T Reaction

5 µl of radio-labelled DNA were mixed with 4 µl of sonicated calf thymus DNA and the sample incubated at 90°C for 2 minutes followed by cooling on ice. 20 µl $KMnO_4$ were added and the reaction incubated at 20°C for 8 minutes. 10 µl of allyl alcohol (ICN Flow) were added and the sample vortexed for 30 seconds to mix. The sample was then dried in a vacuum cetrifuge and the pellet resuspended in 150 µl of ice cold 1 M piperidine. Following incubation at 90°C for 30 minutes, 1.2 ml butan-1-ol were added and the sample vortexed for 30 seconds to mix.

All reactions were then treated in the same way after the steps previously described were performed. The samples were centrifuged at 4°C for 2 minutes at 15 krpm (Sorvall Microspin 24 centrifuge) and the supernatant removed. 150 µl 1% SDS and 1.5 ml butan-1-ol were added and the sample vortexed for 30 seconds. The samples were then centrifuged at room temperature for 2 minutes at 15 krpm (Sorvall Microspin 24 centrifuge). The supernatant was then removed and the samples lypholized for 20 minutes. 10 µl of formamide-EDTA sample loading buffer were added and votexed for 20 seconds. The samples were stored at -20°C until electrophoresis.

Dephosphorylation of DNA

Dephosphorylation of DNA was carried out in a total volume of 50 µl Calf Intestinal Alkaline Phosphatase (CIP) buffer. This was supplied as a 10x stock of which 0.1 volumes were added to an aqueous solution containing 1 µg DNA (~0.625 pmol 5'

terminii). 1 unit of CIP (Boehringer Mannheim) was then added and the reaction incubated at 37°C for 30 minutes. A second aliquot (1 unit) of CIP was added and the incubation continued at 37°C for a further 30 minutes. The reaction was terminated by the addition of 1 µl 0.5 M EDTA and heating to 75°C for 10 minutes. The DNA was purified by phenol-chloroform and chloroform extraction, and ethanol precipitated using sodium acetate (pH7.0). The DNA was resuspended in 1x TE.

DNA Ligation

DNA ligation was caried out in a total volume of 10-20 µl of bacteriophage T4 DNA ligase buffer. This was supplied as a 10x stock of which 0.1 volumes were added to the DNA solution. 40 units of bacteriophage T4 ligase (New England Biolabs) were then added and the reaction incubated overnight at 15°C. The reaction was terminated by heating to 70°C; the solution was allowed to cool to room temperature slowly to promote the reassociation of double stranded DNA. DNA fragments were generally ligated with the vector molecule at a 3-5 fold molar ratio of cohesive ends.

PCR

PCR was carried out directly on bacterial colonies rather than purified DNA. A small amount of the colony of interest was resuspended in 50 µl of sterile distilled water in an Eppendorf tube and boiled for 2 minutes. The tube was then centrifuged for 2 minutes at 15 krpm (Sorvall Microspin 24 Centrifuge). 10 µl of the lysate wa then transferred to a 0.2 ml PCR tube. 2.5 µl Tbr polymerase buffer (NBL) was

added along with 11.5 µl sterile distilled water and 1 µl of each primer used (0.4-1 µM of each). The reaction mix was then layered with PCR wax and the mix stored on ice. An upper layer of 0.5 units Tbr polymerase (NBL), 10 µl dNTP's (250 µM of each), 2.5 µl polymerase buffer and 12.3 µl sterile distilled water was aliquoted on top of the PCR wax to allow hot-start PCR to be performed so as to increase the specificity of the reaction. The conditions for PCR were 30 cycles with 45 seconds at 94°C, 1 minute at 55°C, 1 minute at 72°C plus 5 seconds of extension followed by 10 minutes at 72°C. 5 µl aliquots of the reaction were checked on a 2% Nusieve agarose gel (FMC Bioproducts) prepared in the same way as normal agarose gels (described later).

Gel Electrophoresis

Agarose Gel Electrophoresis

Agarose gels were made and run in 1x TAE gel buffer. The concentration of agarose was a function of the size of the relevant DNA fragments and varied from 0.5-2%. An agarose concentration of 0.8% was generally used to resolve fragments of divergent sizes.

To prepare an agarose gel, the appropriate amount of Seakem agarose (FMC Bioproducts) was dissolved in 100 ml 1x TAE by simmering in a microwave oven. The molten agarose was cooled to 46°C, poured into a 14 cm x 11 cm perspex mould with a 14 tooth comb and allowed to set. The gel was then placed into a BRL

model H5 horizontal gel electrophoresis tank containing 1 litre 1xTAE buffer. DNA samples were prepared by the addition of 0.2 volumes of TAE gel-loading sample buffer and loaded into the the gel slots. Gels were generally run overnight at ~1 V cm^{-1} (~26 V) followed by staining in 0.5 μg ml^{-1} of ethidium bromide for 30 minutes. Following destaining to remove excess ethidium bromide the gel was examined using 304 nm UV light on a C-62 Blak-Ray transilluminator (Ultraviolet Products Incorporated). Photographs were taken on Polaroid 667 or 665 film using a Wratten 25 red filter with an exposure time of 1 second and an aperture of f/5.61/2. DNA fragments for purification were excised using a sterile scalpel blade.

Chloroquine Gel Electrophoresis

To prepare the agarose gel 4 g Seakem agarose (FMC Bioproducts) were dissolved in 400 ml of 1x TBE (10 mM EDTA) by simmering in a microwave oven. The molten agarose was allowed to cool to 46°C and chloroquine added to a final concentration of either 1.8 μg ml^{-1} or 20 μg ml^{-1} (these values were determined experimentally to give the greatest resolution of topoisomers). The agarose was poured into a perspex mould 25 cm x 30 cm, allowed to set. It was then placed into a BRL H2 horizontal gel electrophoresis tank containing 2 litres of 1x TBE buffer (10mM EDTA). Gels were run overnight at ~3 V cm^{-1} (~75 V). The gels was then stained in 0.5 μg ml^{-1} ethidium bromide for 1 hour followed by destaining with repeated washes in distilled water for at least 1 hour. The gel was then photographed as previously described except that an exposure time of 45 seconds used.

Polyacrylamide Gel Electrophoresis

Polyacrylamide gels (40 cm x 21 cm x 0.4 cm) were made and run in TBE gel buffer using a Sequi-Gen Nucleic Acid Sequencing Cell (BioRad). The apparatus was prepared according to the manufacturers instructions.

12% Denaturing Polyarylamide Gels

The glass plates were assembled using 0.4 mm spacers and clamps provided by the manufacturer (BioRad). 60 µl TEMED and 150 µl freshly prepared 10% AMPS were added to 10 ml of 0.5x TBE 12% acrylamide solution and the mixture was promptly used to impregnate a 25 cm x 5 cm strip of blotting paper in a casting tray. The bottom edge of the glass plate sandwich was pushed firmly against the blotting paper and the catalyzed acrylamide was allowed to enter the mould by capillary action. After the acrylamide had set (~2 minutes), thereby sealing the bottom edge, the glass plate sandwich was laid at an angle of ~20° to the horizontal. 40 µl TEMED and 100 µl 10% AMPS were added to 35 ml 0.5x TBE 12% acrylamide solution and the mixture was poured into the gel mould using a 25 ml glass pipette. A 24 well sharkstooth comb was inserted with the flush side in contact with the acrylamide. The top of the mould was covered with Saran wrap and the gel left to set overnight at 4°C. The gel mould was assembled in the apparatus according to the manufacturers instructions. 350 ml 1x TBE buffer were poured into the lower buffer chamber, the gel was clamped into place and the upper buffer chamber filled witth 1x TBE buffer. The comb was then removed and the gel preheated to a temperature of 50-55°C using a constant power of 40 W. DNA samples were denatured by boiling in

a waterbath for 3 minutes followed by rapid cooling on ice. The samples were loaded onto the gel using a micropipetter and the gel was run at constant power of 38 W. After electrophoresis, the gel mould was dismantled and the glass plate with polyacrylamide gel attached placed in 2.5 litres of gel fix for 30 minutes. The gel was transferred to wet blotting paper (Ford Goldmedal), covered in Saran wrap and dried in a BioRad Model 583 Gel Drier for 1 hour at 80°C.

DNA Sequencing Gels

DNA sequencing gels were made up using 6% LongRanger gel solution and run in 1x TBE buffer. The gels were poured as described previously. The DNA sequencing reactions were denatured by boiling for 3 minutes followed by rapid cooling on ice. 2.5 µl of each reaction were then loaded onto the gel in the order ddG, ddA, ddT, ddC. The gel was run at a constant power of 38 W for 2-5 hours. After electrophoresis the gel was transferred onto wet filter paper (Ford Goldmedal) and dried without fixing as described above.

Autoradiography

^{32}P Isotopes

Autoradiography was carried out at -70°C in Cronex (DuPont) cassettes containing "xtra-life" intensifying screens, using Cronex 4 (DuPont)X-ray film (30 cm x 40 cm) preflashed to an OD_{540} of 0.1. The exposure time was varied from 4-16 hours, and films were developed in an X-OGRAPH Compact X2 automatic film processor.

^{35}S Isotopes

Autoradiography was carried out at room temperature in Cronex (DuPont) cassettes, using Cronex 4 (DuPont) X-ray film (30 cm x 40 cm). The exposure time varied for 1-3 days and the films were developed in an X-OGRAPH Compact X2 automatic film processor.

RESULTS

CHAPTER 3

INSTABILITY OF DNA PALINDROMES IN PLASMID REPLICONS

Introduction

Many of the spontaneous deletions which occur in general sequence DNA are comparatively rare events, for example in the *lac* operon where they are detected at a frequency of between 10^{-8} and 10^{-7} (Farabaugh *et al.*, 1978). The presence of palindromic DNA has been shown to increase the frequency of deletion formation mediated by illegitimate recombination (see chapter 1) and as a result using palindrome instability as an example of an illegitimate event provides greater opportunities for investigating deletion formation. Many of the studies performed to date have involved the use of relatively short palindromic sequences (DasGupta *et al.*, 1987; Weston-Hafer and Berg, 1989; Kazic and Berg, 1991) and although these palindromes are longer than any encountered naturally in *E. coli* they are short in comparison to those found in higher organisms.

An investigation was made into the deletion of long palindromic sequences (109 bp and 571 bp) from high copy number pUC based plasmids. In particularly the effect of local DNA sequence, i.e. the presence of direct repeats; the influence of the direction of replication through the region of plasmid containing the palindrome and the effect of different *E. coli* mutations on the propagation of plasmids bearing long palindromes.

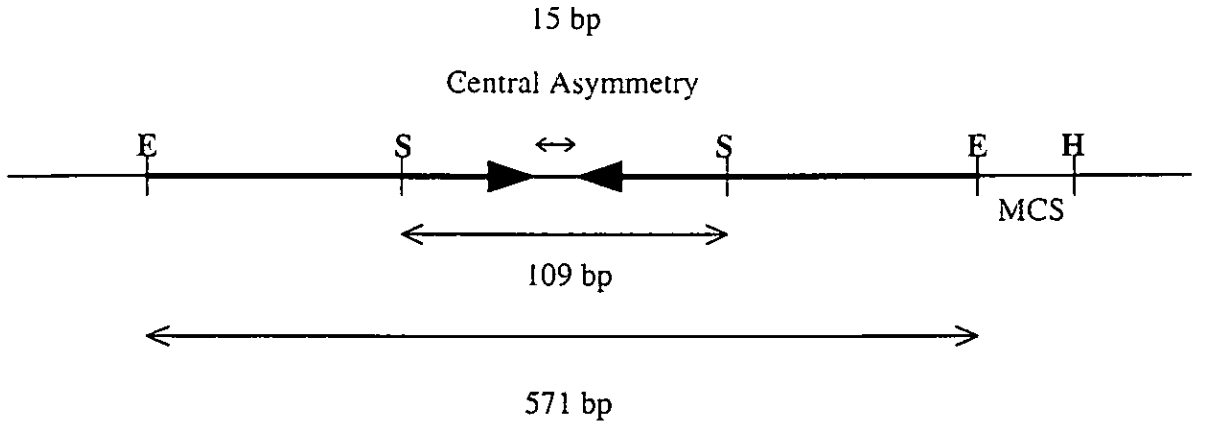
Results

Factors Affecting the Stability of a Palindrome in Plasmid Replication

The system used was based on a 571 bp near perfect palindrome cloned into the *Eco*RI site of the high copy number plasmid pUC18 (Chalker, Shaw and Leach, unpublished). This plasmid was denoted pAC2. The palindrome was itself a deletion product of a 3.2 kb artificially created palindrome (Leach and Stahl, 1983). It contained a central region of asymmetry that was 15 bp in length and there were two internal *Sac* I sites that defined a central fragment of 109 bp (Leach *et al.*, 1989) (see figure 5). The *Eco*RI site used in the cloning was present within the multiple cloning site of the plasmid and insertion of the palindrome into this site resulted in insertional inactivation of the *lacZ* gene and so a lack of β-galactosidase activity.

When the *E. coli sbcC* strain DL494 was transformed with the plasmid containing the palindrome and the cells plated out on medium containing the chromogenic substrate for β-galactosidase (X-gal), the majority of the colonies had a sectoring phenotype (see figure 6). Extraction of the plasmid DNA from *E. coli*, followed by digestion with a restriction enzyme that cut the plasmid only once (e.g. *Hind*III) resulted in two distinct bands when the DNA was electrophoresed on an agarose gel (see figure 7). The upper band had the predicted size of linearised

Figure 5. Structure of 571 bp palindrome in pAC2



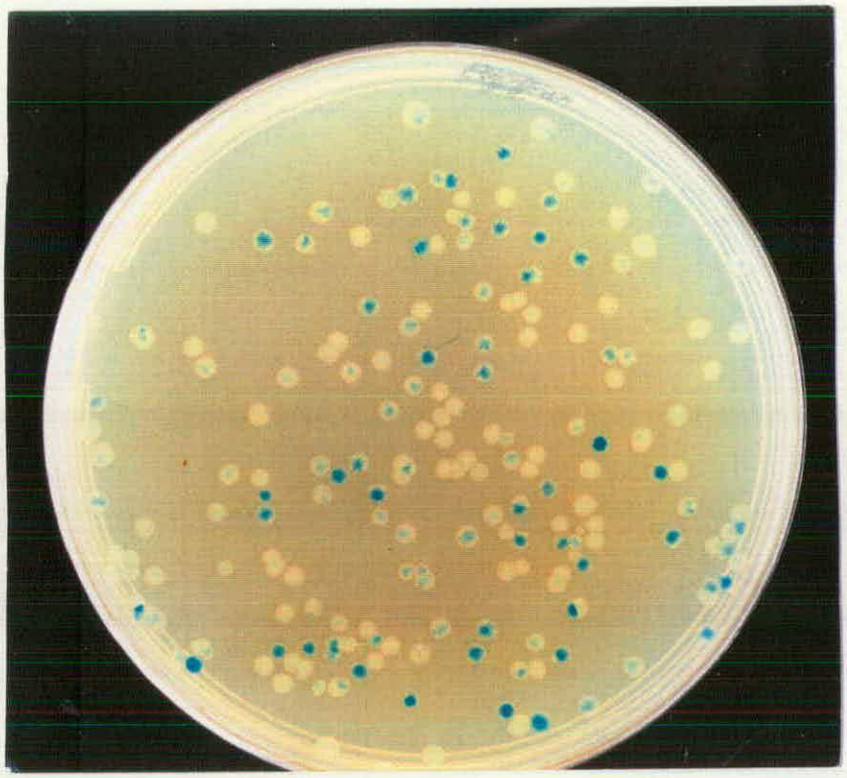
N.B Not drawn to scale

H, S, E Restriction enzyme sites (*Hind*III, *Sac*I and *Eco*RI)

MCS Multiple cloning site of the plasmid

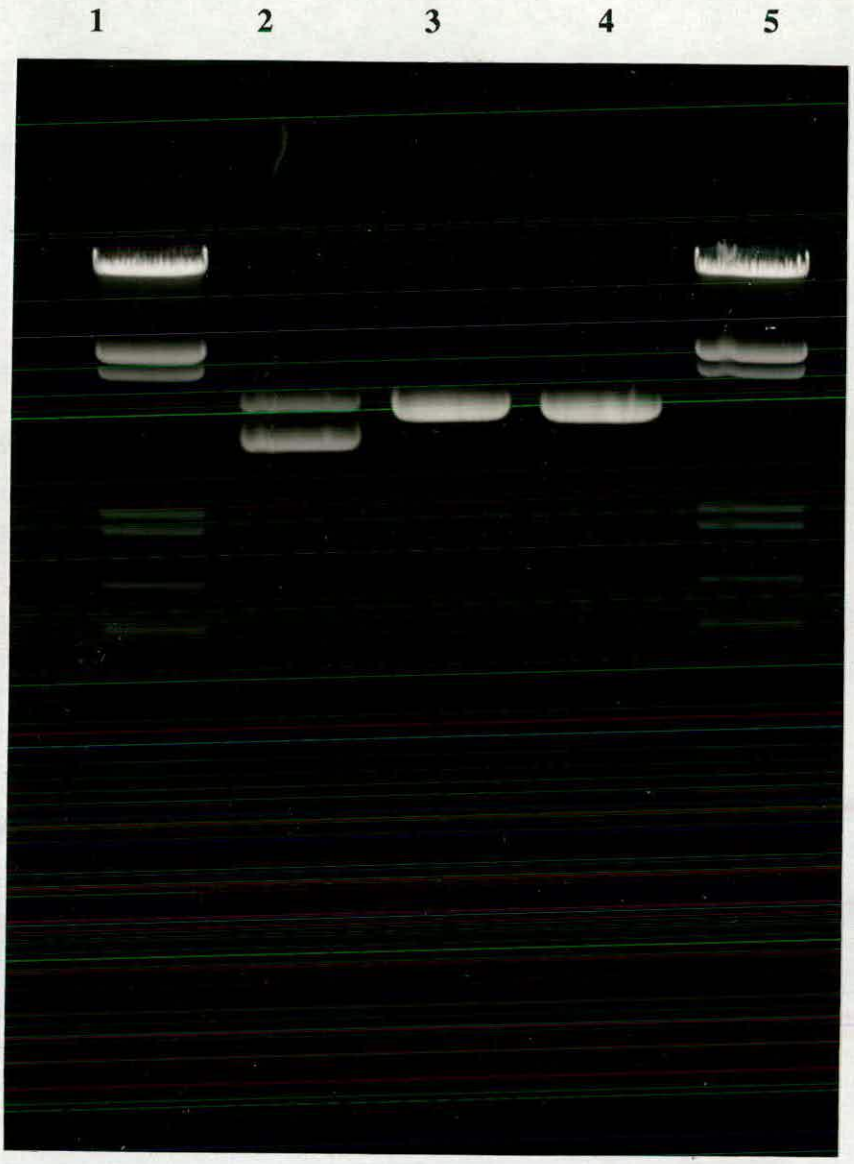
The palindrome in pAC2 is 571 bp in length, with a central asymmetric region of 15 bp. The palindrome is flanked by *Eco*RI restriction sites. Two *Sac*I restriction sites define a smaller internal palindrome of 109 bp length.

Figure 6. Strain containing plasmid pAC2 plated on medium containing ampicillin and X-gal.



Strain DL324 plated on selective medium containing ampicillin and the chromogenic substrate X-gal. The blue colour represents those cells bearing plasmid which is producing a functional β-galactosidase.

Figure 7. *Hind* III digest of plasmid DNA



Key

1. Lambda marker *Eco*RI/*Hind*III digest
2. pAC2
3. pMS5
4. pMS7
5. Lambda marker *Eco*RI/*Hind*III digest

plasmid containing the palindrome. The lower band was of a size consistent with deletion of the palindrome. It was apparent that there was a preferred deletion product since only one band of plasmid DNA without palindrome was seen. Approximately 50% of the plasmid DNA appeared in a deleted form.

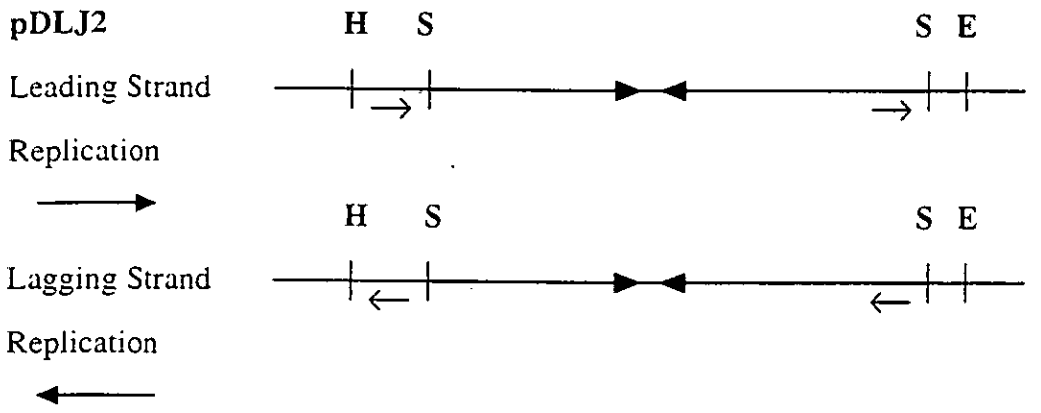
The 571 bp palindrome contained an internal 109 bp palindrome bounded by the *Sac*I sites. The 109 bp palindrome was subcloned into the related plasmids pUC118 and pUC119 to give plasmids designated pDLJ1 and pDLJ2 respectively (D.Leach, unpublished). The plasmids differed with respect to the orientation of the multiple cloning sites present within them. When an *sbcCD* strain containing the plasmids was plated out on medium containing X-gal and IPTG the colonies appeared to have flecks of blue colour rather than the blue sectors seen with cells containing pAC2. A higher deletion frequency (i.e. degree of blue flecking) was seen for pDLJ2 than pDLJ1 (data not shown). The deletion frequency was apparently much lower than that seen for plasmid pAC2 and only a single band was seen upon digestion of the plasmid with *Hind*III. Blue colonies were isolated and purified by restreaking to obtain plasmid which did not carry the palindrome. When plasmid DNA was isolated from blue colonies and the multiple cloning site sequenced a seven base pair direct (5'-TTGCATG-3') repeat was identified as the deletion endpoint. One of the repeats was located within the multiple cloning site of the plasmid downstream of the *Hind*III restriction site and the other repeat was present within the palindrome (figure 8). Deletion of the palindrome via the direct repeats

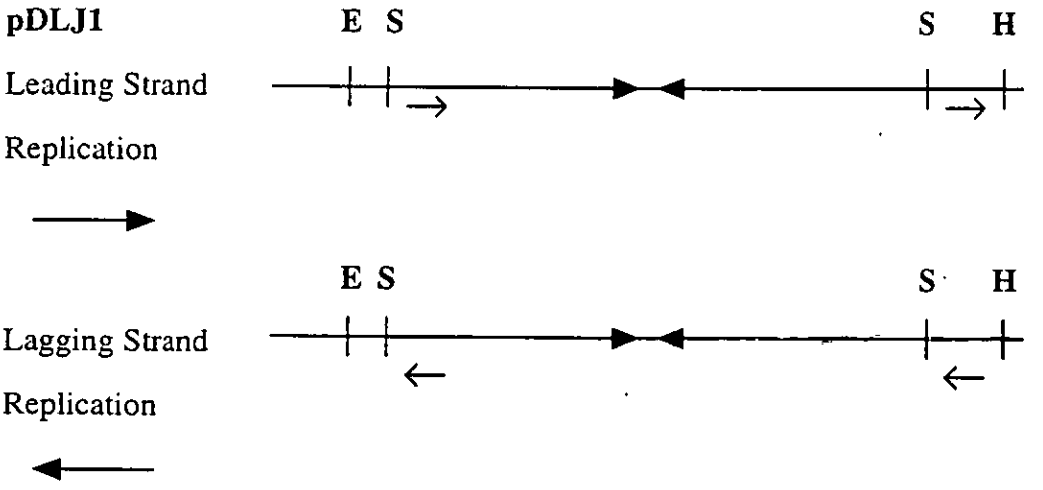
resulted in 4 bases (5'-TCCA-3') of the palindromic sequence being retained in the plasmid (data not shown). The preference for one of the repeats to be located just inside the palindrome and one to be located just outside the palindrome was consistent with model in which replication proceeded some way into the palindrome before stalling.

The asymmetry of the direct repeats allowed testing of the hypothesis of whether deletion of the palindrome occurred preferentially on the leading or lagging strand of the replication fork. Inversion of the fragment of DNA that included the palindrome and direct repeats in plasmids pUC118 and pUC119 moved the preferred orientation of the repeats between the leading and lagging strands (figure 8). As a replication fork entered the palindrome in one direction, strand slippage would be favoured if a good target repeat was located downstream of the palindrome. A good target sequence is one which is located just outside the palindrome and shares several bases complementarity with a donor sequence present just within the palindromic sequence.

The highest frequency of palindrome deletion was obtained when inversion of the DNA fragment containing the palindrome and the repeat placed the correct orientation of repeat on the lagging strand of the replication fork. Experiments by Trinh and Sinden (1991) suggested that slippage occurred more frequently on the

Figure 8. Position of direct repeats on leading and lagging strands of replication fork





H, S, E Restriction enzyme sites (*Hind*III, *Sac*I and *Eco*RI)

⟶ Direct repeat drawn with respect to the direction of replication

⟵ Direction of replication fork

lagging strand of the replication fork. However, similar experiments by Weston-Hafer and Berg (1991) showed no difference between deletion on the two strands. This work confirms the former observation using a different system and adds weight to the conclusion that there can be a bias for slippage on the lagging strand of the replication fork. The hypothesis was further supported by the fact that the direct repeats that flank the palindromes, present as a direct consequence of cloning the palindrome, were never used as deletion endpoints by the 109 bp palindrome.

Plasmid Derivatives with Increased Palindrome Stability

A series of plasmids were isolated that showed a decreased level of palindrome instability (Shaw and Leach, unpublished). The plasmids were isolated from cells containing the plasmid pAC2. Two of the plasmids, pMS5 and pMS7, were investigated to attempt to determine the source of the increased stability. When plated on media containing X-gal and ampicillin, pMS5 exhibited a deletion frequency of ~2% and pMS7 a deletion frequency of <0.002% as determined by transformation efficiency into a wild-type and an *sbcCD* strain (tables 3 and 4). Few blue colonies were detected with these plasmids (see figures 9 and 10). When DNA was prepared from cells containing the plasmids, restricted with the enzyme *Hind* III and electrophoresed on an agarose gel (see figure 7), a major band was seen for pMS5 with a faint band equivalent to a deletion product running just ahead of the

major band. Only a single band was seen for pMS7 consistent with the very low deletion frequency determined for this palindrome.

Purification of the rare blue colonies produced by cells containing these plasmids followed by sequencing of the multiple cloning sites identified that a three base pair direct repeat (5'-TGA-3') was used as a deletion endpoint. One of the repeats was present just outside the multiple cloning site of the plasmid and the other repeat was within the palindrome itself. Deletion of the palindrome resulted in three bases (5'-AAT-3') of the palindrome sequence remaining within the plasmid (see figures 11 and 13). The original sequence of pUC18 is shown in figure 12 with the *Eco*RI restriction site used in the cloning of the palindrome highlighted. Figure 13 shows the sequence of pMS7 after deletion of the palindrome via a three base direct repeat. The three bases of palindrome sequence remaining in the plasmid are also shown in figure 13.

Re-introduction of the plasmids into *E. coli sbcC* or *sbcCD* strains (DL494 or DL733) resulted in the phenotype being transferred with the plasmid. This was consistent with the source of the increased stabiliity being present on the plasmid and not within the genome of the original *E. coli* cell. As a result a number of molecular techniques were employed to determine whether the increased stability was the result

Table 3. Number of transformants obtained after plating on selective medium.

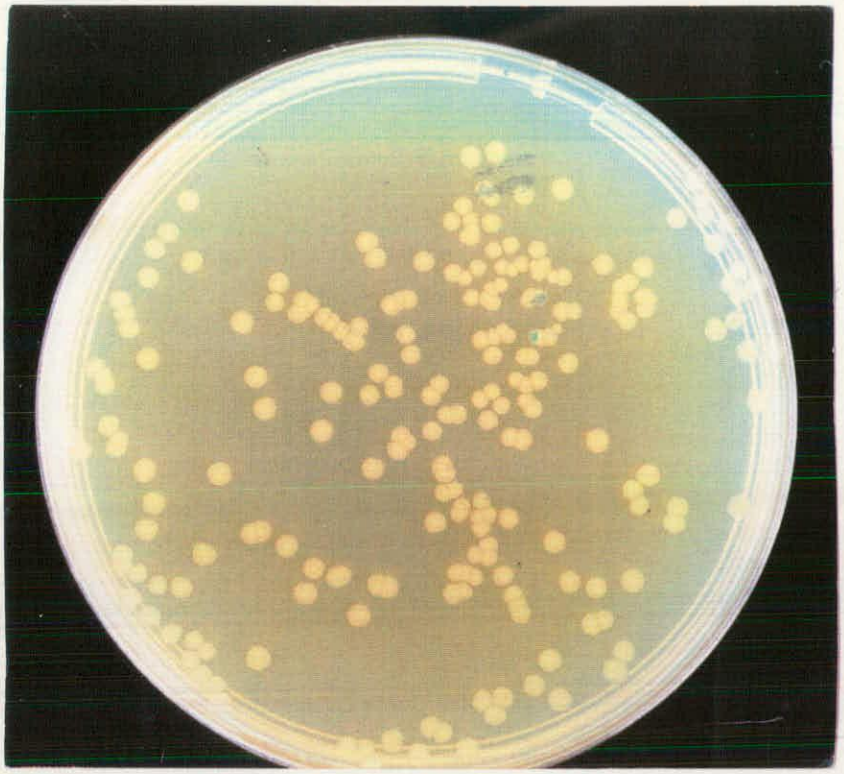
Strain	Plasmid		
	pAC2	pMS5	pMS7
DL324 (wt)	727	23	<1
DL733 (Δ*sbcCD*)	1349	866	551

The number of colonies arethose obtained after transformation of a wild-type and *sbcCD* strain with 100 ng of plasmid DNA followed by plating on selective medium.

Table 4. Transformation efficiency compared in wild-type and *sbcCD* backgrounds

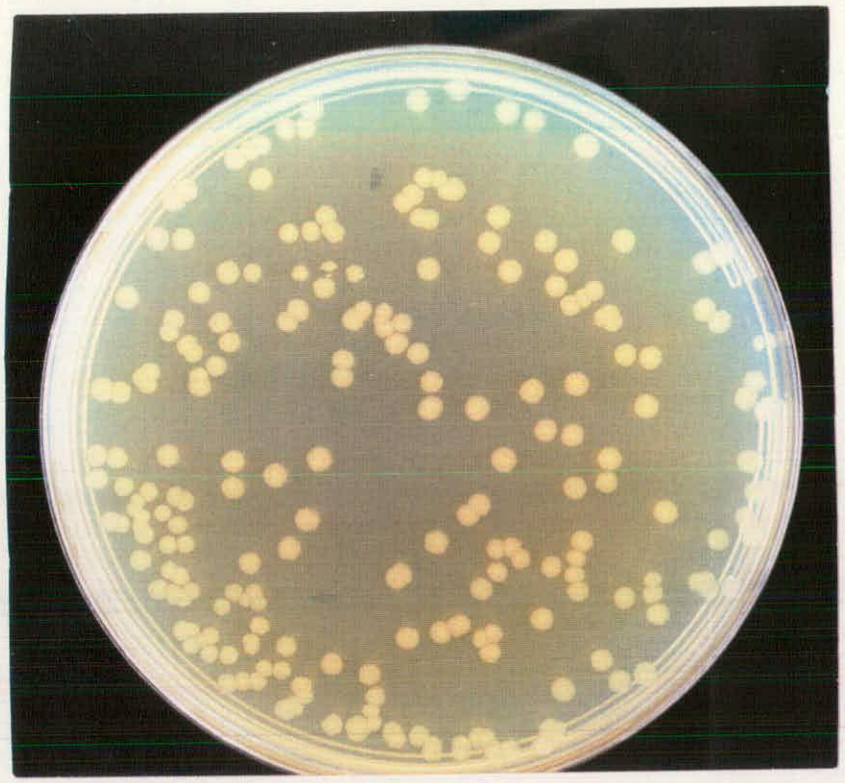
Plasmid	No. of wild-type transformants/ No. of *sbcCD* transformants
pAC2	727/1349=0.540
pMS5	23/866=0.027
pMS7	<1/551=<0.002

Figure 9. Strain bearing pMS5 plated on medium containing ampicillin and X-gal.



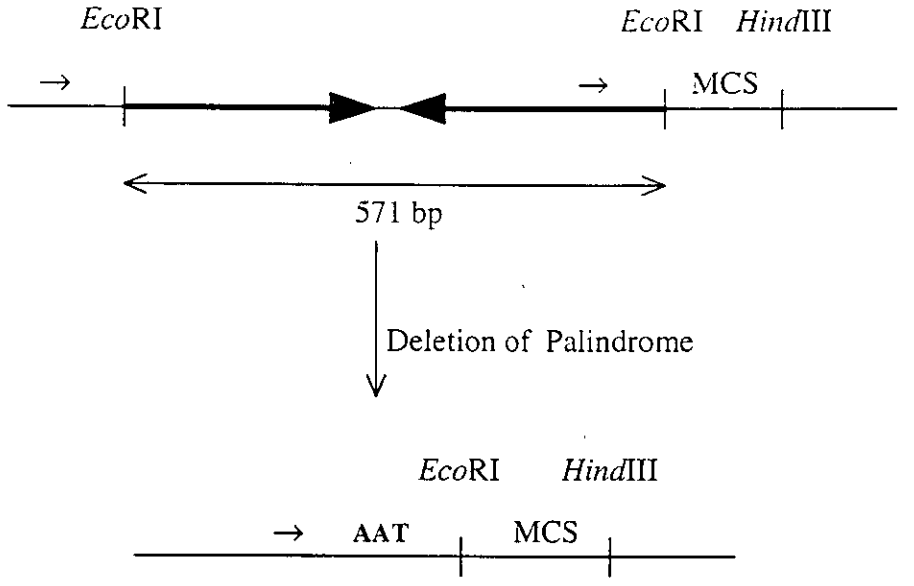
Strain bearing plasmid pMS5 plated on selective medium containing ampicillin and the chromogenic substrate X-gal. The blue colour represents those cells bearing plasmid which is producing a functional β-galactosidase.

Figure 10. Strain bearing pMS7 plated on medium containing ampicillin and X-gal.



Strain bearing plasmid pMS7 plated on selective medium containing ampicillin and the chromogenic substrate X-gal. The blue colour represents those cells bearing plasmid which is producing a functional β-galactosidase.

Figure 11. Position of direct repeats used as deletion endpoints in pMS7.



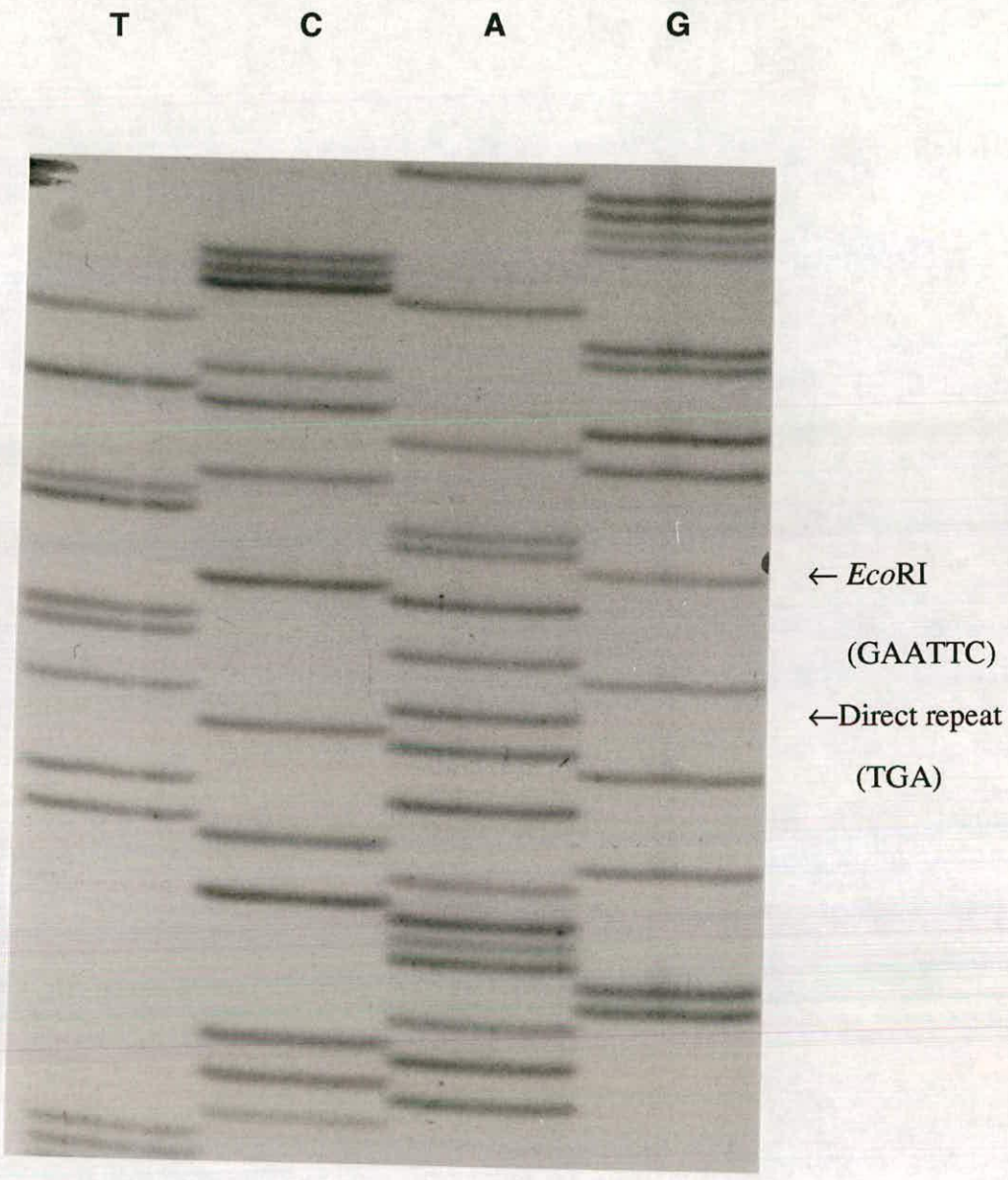
N.B Not drawn to scale.

→ TGA direct repeat used in deletion

AAT Region of palindromic sequence remaining after deletion

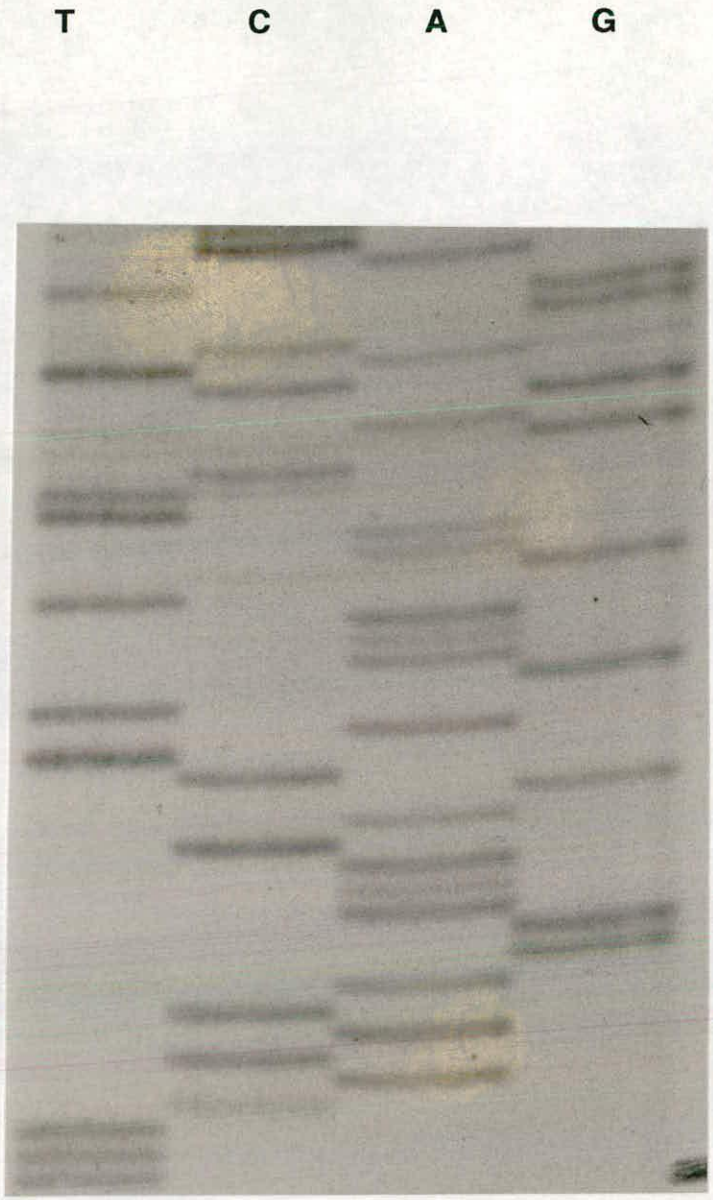
MCS Multiple cloning site

Figure 12. Autoradiograph of pUC18 multiple cloning site



Sequence produced by dideoxy chain termination sequencing. *Eco*RI restriction site used in the cloning of the palindrome is highlighted along with the three base direct repeat used in deletion of the palindrome.

Figure 13. Autoradiograph of deleted form of plasmid pMS7



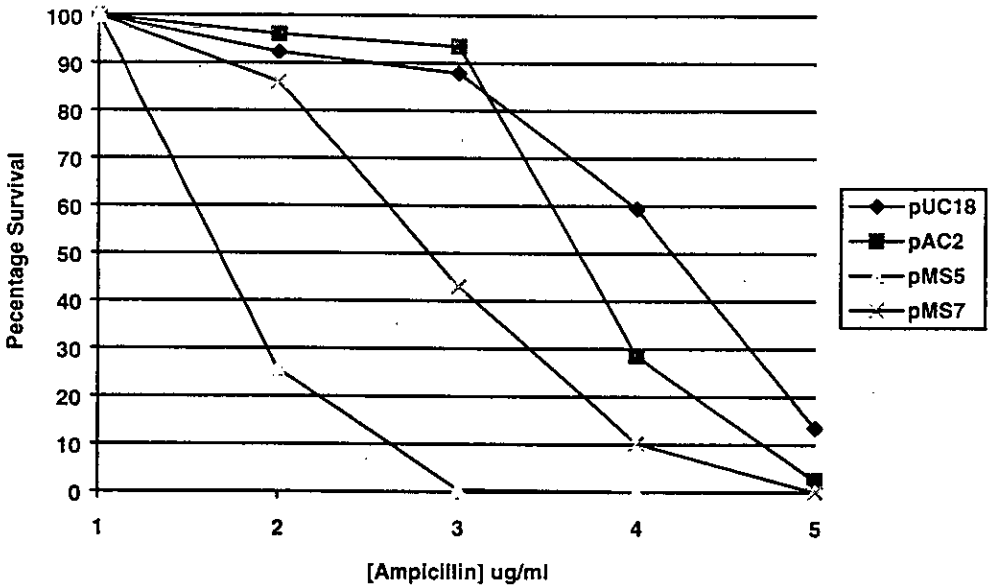
Sequence produced using dideoxy chain termination sequencing. *Eco*RI site used in cloning of the palindrome. AAT is sequence of three bases of palindrome remaining in the plasmid after deletion. TGA sequence of direct repeat used in deletion.

of a plasmid-borne mutation. The region around the multiple cloning site was sequenced to determine if the increased stability was due to a mutation in this region. The sequences obtained for pMS5 and pMS7 were compared to those of pAC2 and pUC18. A mutation was identified in the multiple cloning site of the plasmid but this was present in the mutants pMS5, pMS7 along with pAC2 and pUC18 (data not shown). This mutation was present within a number of commercially available pUC18 preparations (Lobet *et al.*, 1989). The palindromes from pMS5 and pMS7 were sequenced by partial restriction of the plasmids with the enzyme *Sac*I. Digestion generated two fragments, one being composed of the left arm of the palindrome along with the central 109 bp palindrome, this was ligated into pUC19 restricted with *Sac*I. The second fragment was made up of the right arm of the palindrome and the backbone of the plasmid. This was self-ligated to re-circularise the plasmid. Sequencing of both these constructs failed to identify a mutation. The copy numbers of the plasmids were approximately estimated to investigate whether the increased stability observed in pMS5 and pMS7 was due to a significant decrease in the copy numbers of the plasmids. The experiment was performed by plating overnight cultures of cells containing the plasmids on media containing increasing concentrations of ampicillin from 1 mg ml^{-1} to 5 mg ml^{-1}. The percentage survival of each of the strains was calculated and an example of the results obtained are shown in table 5 and figure 14. The copy number of pMS7, as determined in the experiment, was higher than that of pMS5 even though pMS7 was known to have a higher degree of palindrome stability.

Table 5. Determination of plasmid copy number.

[Ampicillin]	% Survival of Strain Containing Plasmid			
mg ml^{-1}	pUC18	pAC2	pMS5	pMS7
1.0	100.0	100.0	100.0	100.0
2.0	92.3	96.1	25.6	86.0
3.0	87.9	93.5	0.5	43.0
4.0	59.4	28.6	0	10.3
5.0	13.5	2.5	0	0

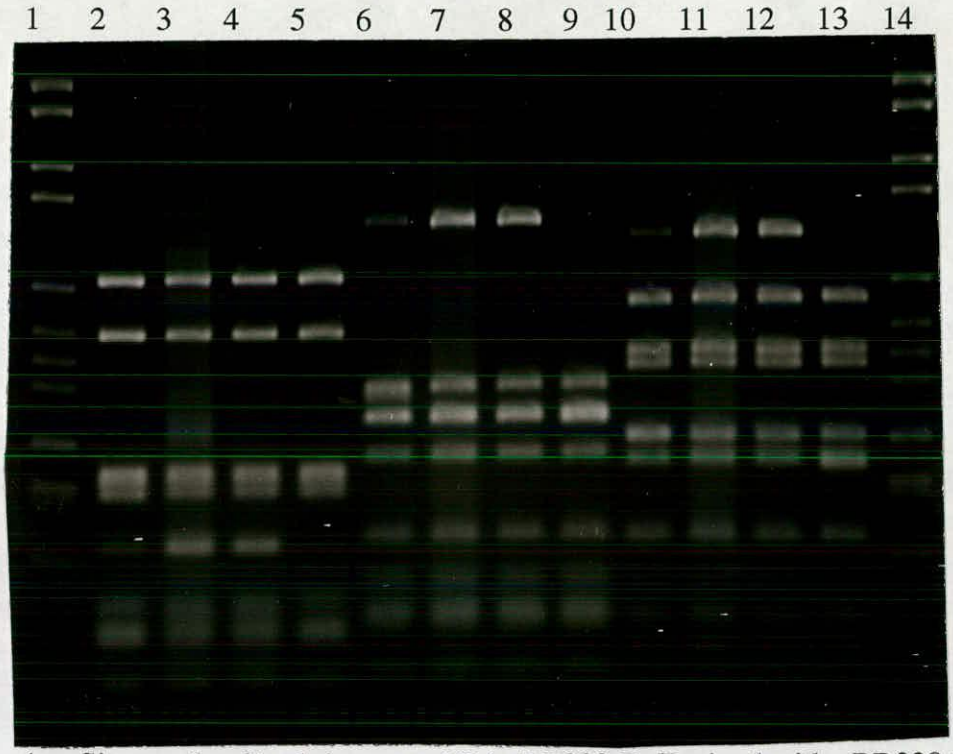
Figure 14. Percentage Survival of Strains Bearing Plasmids on Increasing Concentrations of Ampicillin



A series of restriction fragment swaps was undertaken to identify a region of the plasmid that may contain the mutation responsible for the increased palindrome stability. The experiment involved replacing specific DNA fragments in pMS5 and pMS7 with the corresponding fragment from pAC2 and pUC18. The fragmentsreplaced were defined by the restriction sites *Alw*NI to *Bam*HI and *Afl*III to *Sca*I. It was expected that replacing the region in pMS5 or pMS7 that contained the mutation with the corresponding region from pAC2 or pUC18 would alter the plating behaviour of cells bearing the recombinant plasmid. In other words, the colonies would be blue or sectored instead of white. All DNA fragments were purified from agarose gels prior to ligation using GENECLEAN to minimise the carry over of uncut plasmid DNA. The results obtained were somewhat difficult to interpret but it appeared that the mutation lay in the region defined by the *Alw*NI and *Afl*III restriction sites. Primers were designed that were specific to sites flanking the *Alw*NI and *Afl*III. This region contained the replication origin of the plasmid and it was possible that a mutation in this region could affect the rate of replication and so increase the chance of the palindromes being replicated faithfully. However, sequencing failed to identify any mutation in this region.

The origin of the mutation was further investigated by restriction analysis. This involved digesting the plasmid with a variety of restriction enzymes which have 4 bp recognition sites and therefore have a high frequency of sites in DNA. The enzymes used were *Alu*I, *Hha*I and *Hae*III and following digestion the DNA was

Figure 15. Restriction analysis of plasmids.



1 Size marker Boehringer no.VI (pBR328 *Bgl*I mixed with pBR328 *Hinf*I)

2 pAC2 digested with *Alu*I

3 pMS5 " "

4 pMS7 " "

5 pUC18 " "

6 pAC2 digested with *Hha*I

7 pMS5 " "

8 pMS7 " "

9 pUC18 " "

10 pAC2 digested with *Hae*III

11 pMS5 " "

12 pMS7 " "

13 pUC18 " "

14 Size marker Boehringer no.VI (pBR328 *Bgl*I mixed with pBR328 *Hinf*I)

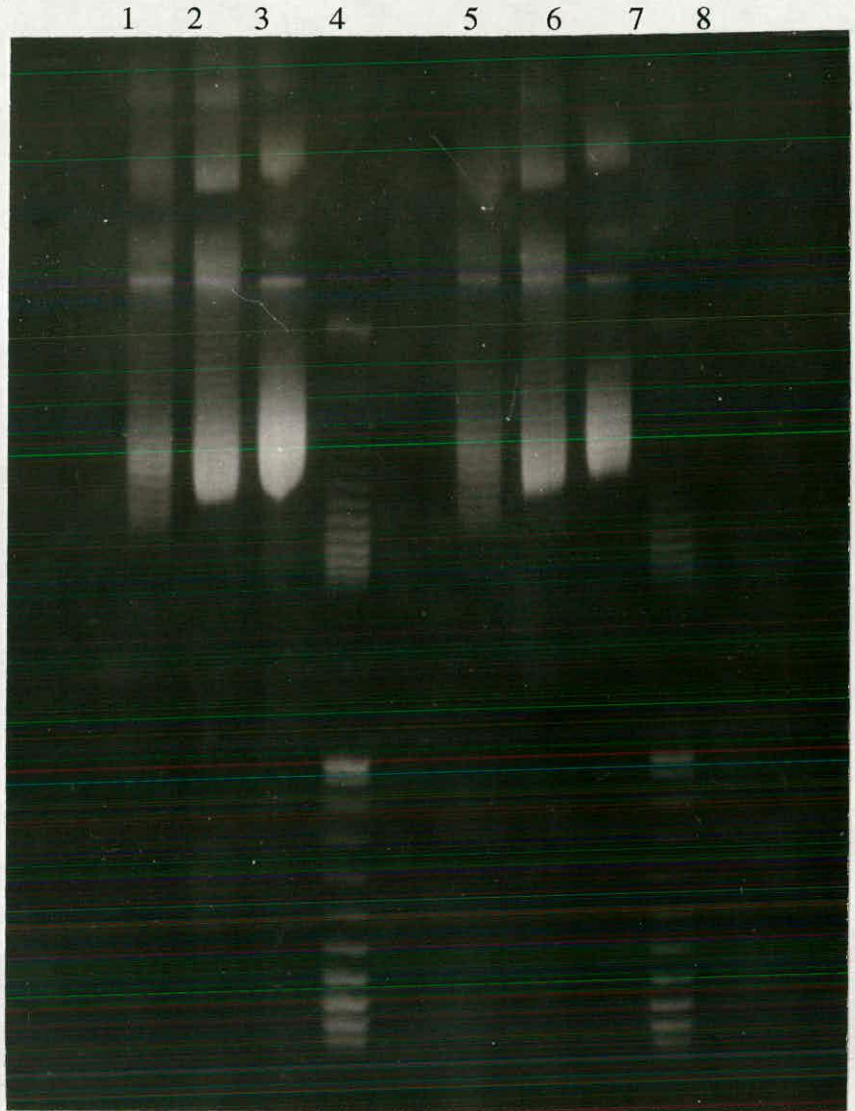
electrophoresed on a 2% Metaphor agarose gel (FMC Bioproducts). pUC18 DNA was also restricted and run on the gel to act as a control. The results are shown in figure 15. There was no difference in the restriction patterns observed for the original pAC2, the mutants pMS5 and pMS7, and the pUC18 control. This indicated that any mutation present did not affect the restriction sites of these enzymes. Neither were there any obvious differences in the sizes of the bands observed.

Primers were designed to produce overlapping sequences that covered the rest of the plasmid backbone. The sequences of pMS5 and pMS7 were compared to that of both pAC2 and pUC18. The only mutations identified were present in both the mutant plasmids as well as pAC2 and pUC18. One region of the palindrome proved impossible to sequence with conventional dideoxy chain termination sequencing. The site, at position 1961-1966 bp, contained a naturally occurring palindromic sequence which proved to be unusually stable. The presence of a mutation within this palindrome might have resulted in it preferentially extruding and so reducing the supercoiling level of the palindrome. This may have prevented supercoiling driven extrusion of the 571 bp palindrome and so decrease its deletion frequency. Maxam-Gilbert sequencing was employed, which involved chemical cleavage of the DNA so the method was not affected by the presence of secondary structures. No differences were visible between the sequences obtained for pMS5, pMS7, pAC2 and pUC18.

One final method was employed to identify any differences between the plasmids involved in the study. The level of supercoiling present within each of the

plasmids was determined by electrophoresing uncut plasmid DNA through agarose gels containing the chemical chloroquine. Chloroquine is an intercalating agent that inserts between the bases of the DNA and so causes relaxation of the DNA. The negative twist change is compensated for by a positive writhe i.e. a negatively supercoiled molecule becomes more positively supercoiled. The increase in positive supercoiling is proportional to the amount of chloroquine added. Two different concentrations of chloroquine were utilised, a low concentration of 2μg ml^{-1} and a high concentration of 20μg ml^{-1}. It was found that the lower concentration gave the most satisfactory results as the higher concentration quenched the fluorescence from ethidium bromide used to visualise the DNA under a U.V. light source. A differencewas observed between pMS5, pMS7, pAC2 and the pUC18 control. When exposed to a low concentration of chloroquine pMS5 and pMS7 migrated behind pAC2 indicating that the plasmids were relatively more relaxed compared to pAC2. pMS7 appeared to be more relaxed than pMS5 (figure 16). It is unknown whether the more relaxed state of pMS7 was a result of a yet unidentified mutation or whether the increased stability of palindromes in pMS7 was a consequence of the lower level of negative supercoiling in this plasmid.

Figure 16. Supercoiling level determination using chloroquine.



1% agarose gel containing 2μg ml^{-1} cloroquine run at 70V, 16 hours

1 & 5. pAC2

2 & 6. pMS5

3 & 7. pMS7

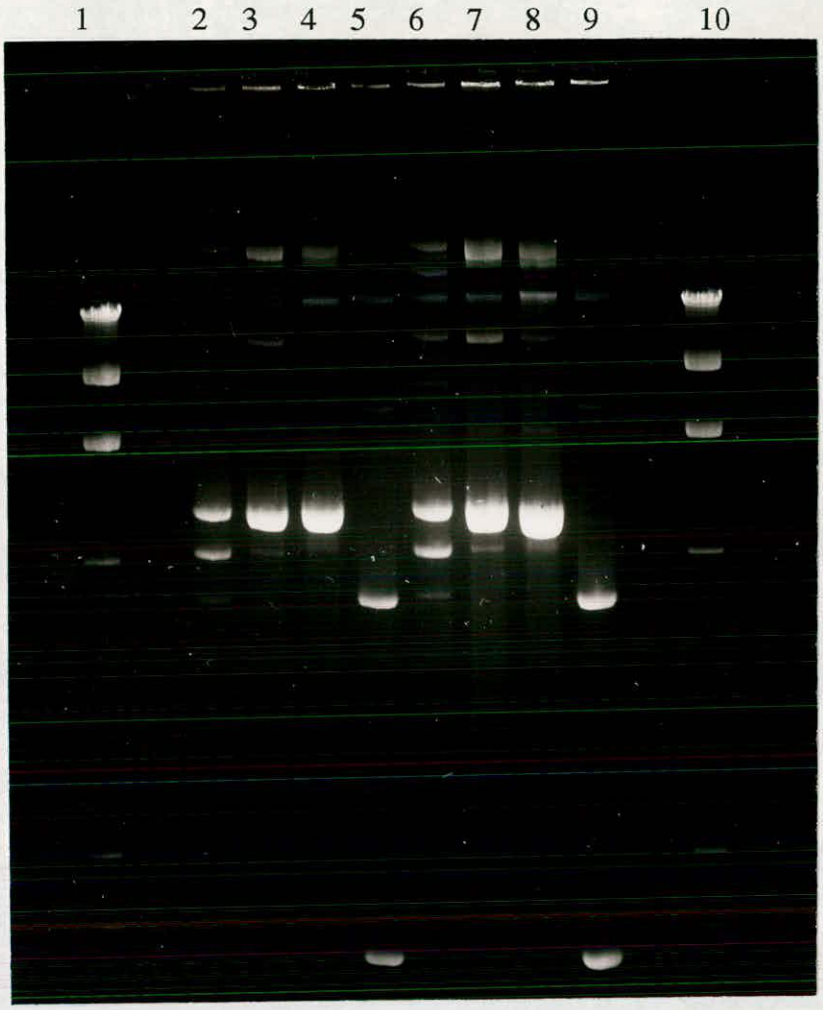
4 & 8. pUC18

Effect of host Genotype

DNA was purified from the *sbcCD* strains bearing plasmids pAC2, pMS5, pMS7 and pUC18 and the DNA electrophoresed uncut on agarose gels to look at the effect of host genotype on the form of the DNA (see figure 17). The plasmid DNA appeared to be highly multimeric. Using densitometry the relative percentages of each of the plasmid multimers were determined. The scans produced by densitometry are shown in figures 18, 19, 20 and 21. Plasmid pAC2 is present predominantly as dimers with 49.8% of the DNA being in this form (peaks 12 and 13). The remainder of the DNA is present in higher multimeric forms, no monomers were visiblePlasmid pMS5 is also present mainly as dimers, 50.9% (peaks 9 and 10), but some monomeric DNA is present (peak 13, 3.7%). Plasmid pMS7 is only present as supercoiled multimers, with dimers predominating (peak 11, 55.7%). Plasmid pUC18 is present in all multimeric forms with monomers accounting for 29.5% of the total DNA content (peak 11).

To examine the effect of a *recA sbcCD* background on the deletion of the 571 bp palindromes in plasmids pAC2 and pMS7, monomeric DNA was artificiallycreated by restricting the DNA with an enzyme that cut only once, followed by self-ligation to recircularise the plasmid. A *recA sbcCD* host (DL888) was then transformed and the cells plated out on medium containing ampicillin, X-gal and IPTG. From the transformations only one white colony was obtained for

Figure 17. Distribution of plasmid multimers in an *sbcCD* host.



Lanes.

1 & 10. Lambda *Hind*III size marker

2 & 6. pAC2

3 & 7. pMS5

4 & 8. pMS7

5 & 9. pUC18

Figure 18. Distribution of plasmid multimers in pAC2.

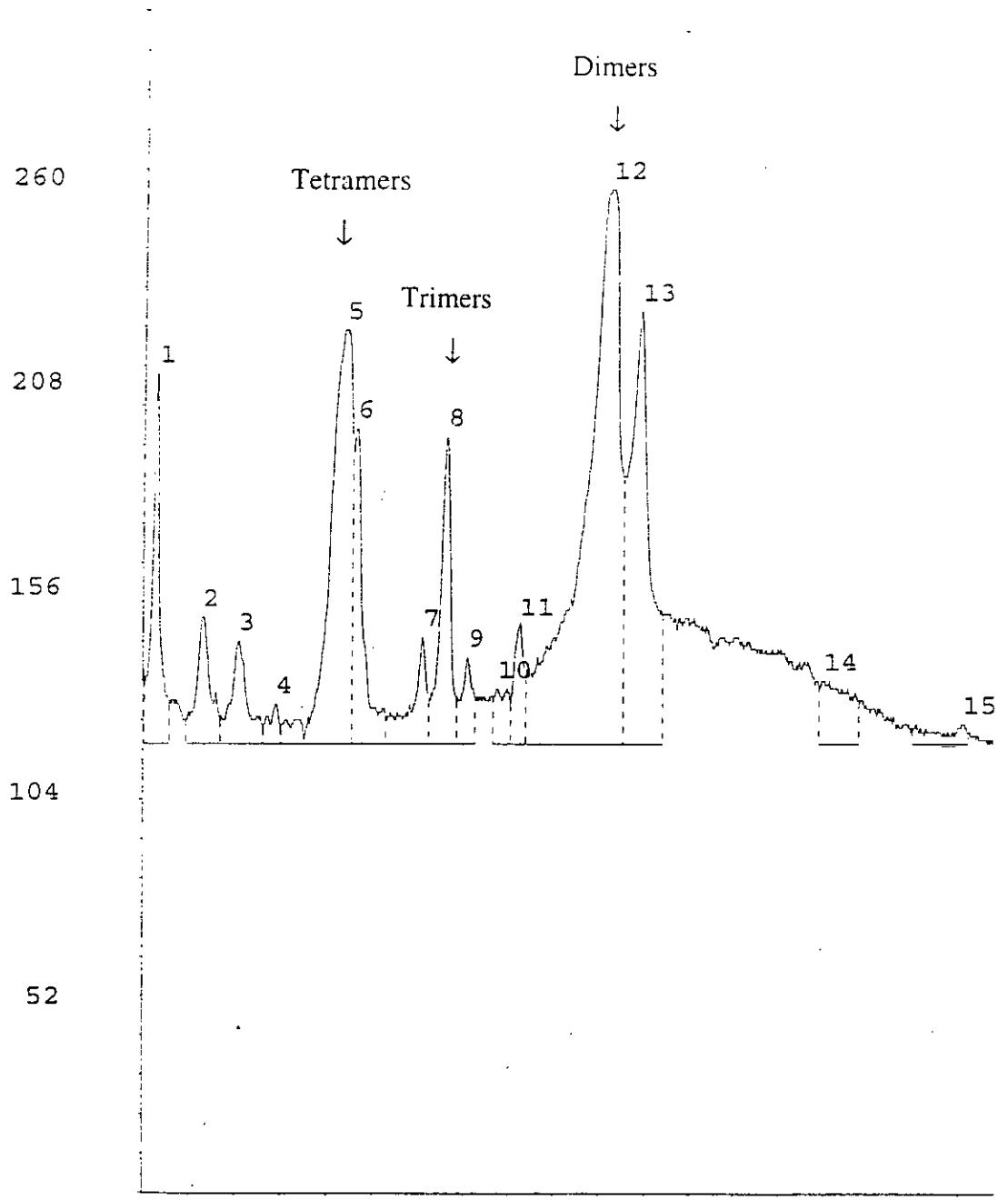


Figure 19. Distribution of plasmid multimers in pMS5

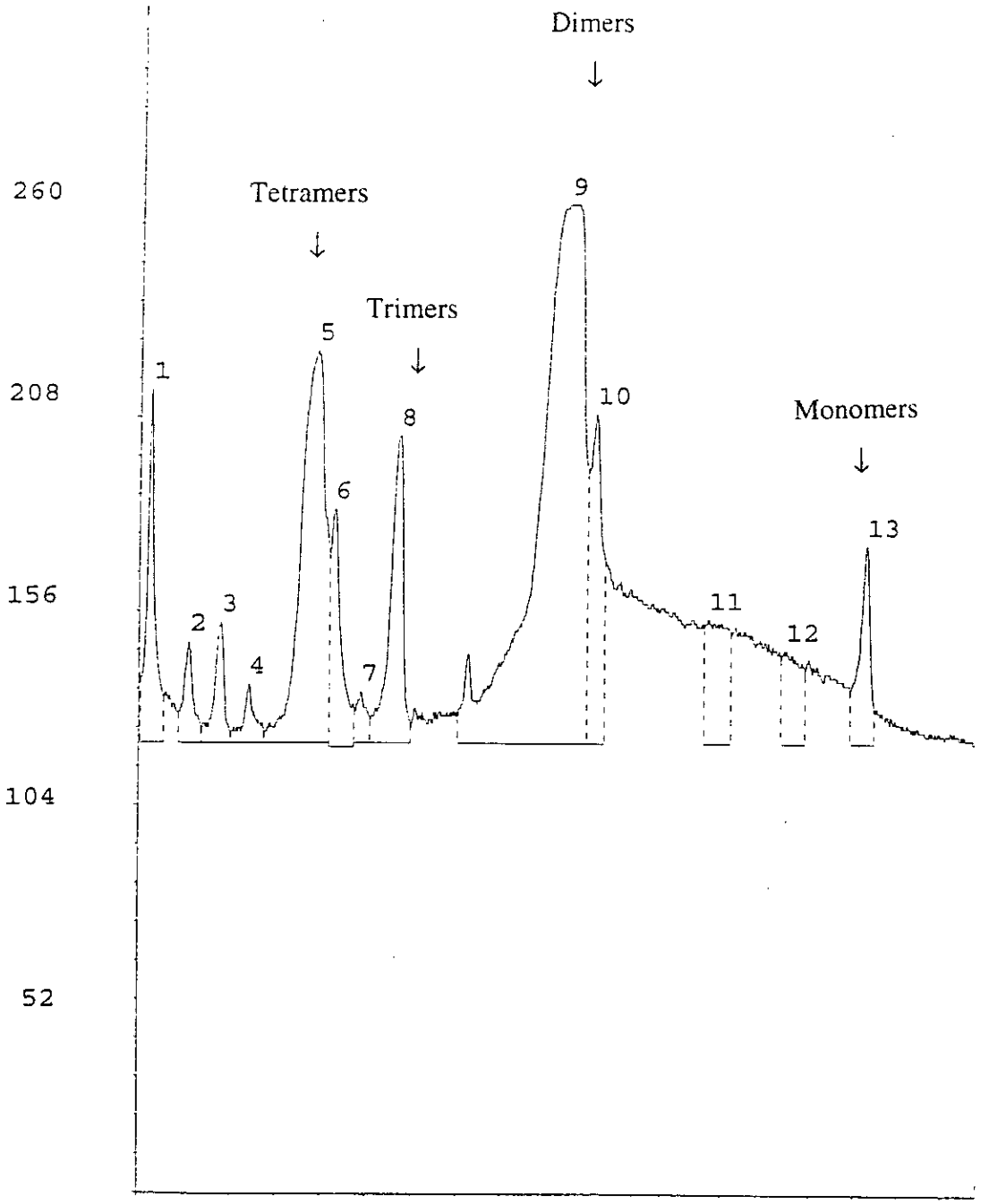


Figure 20. Distribution of plasmid multimers in pMS7

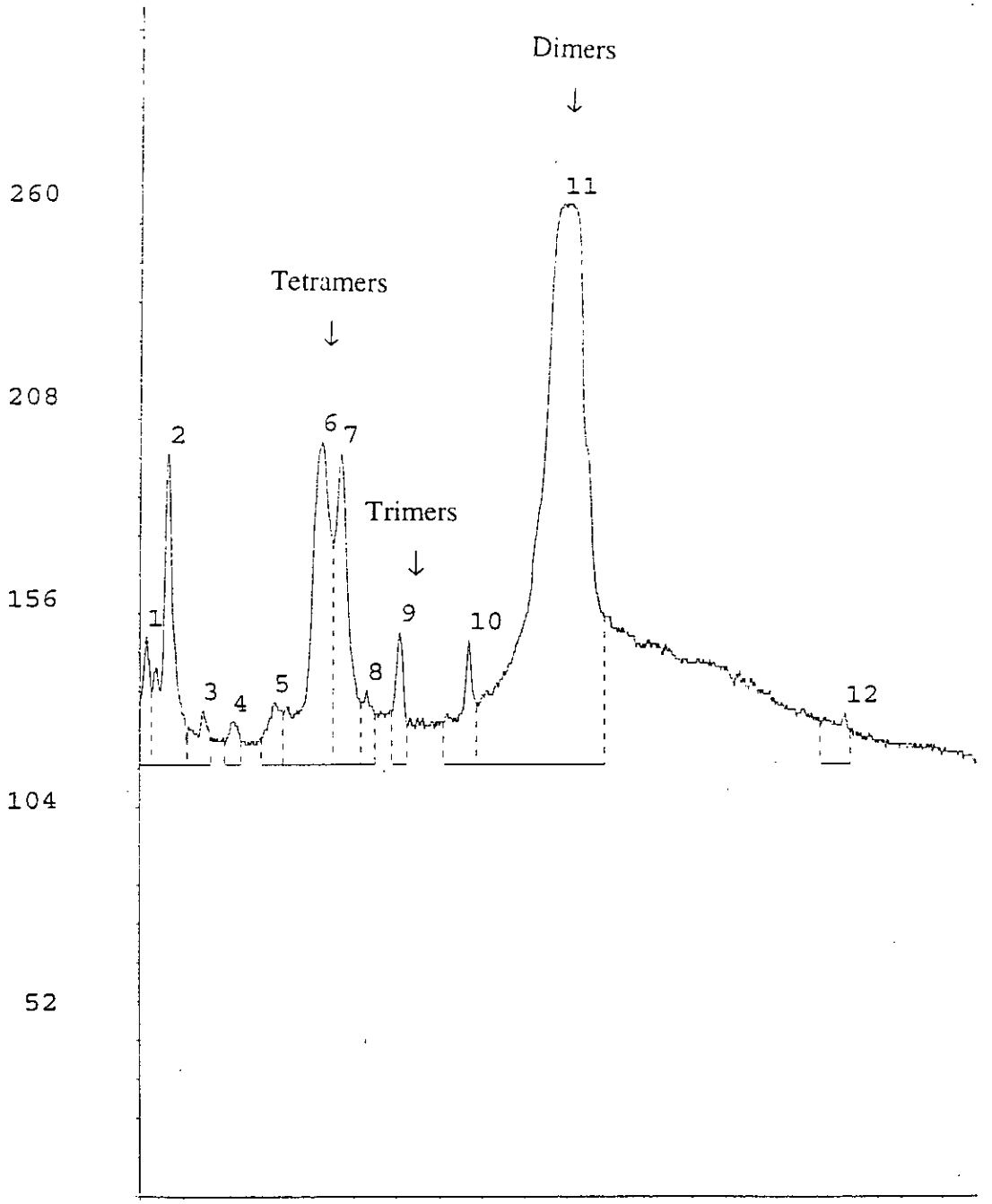
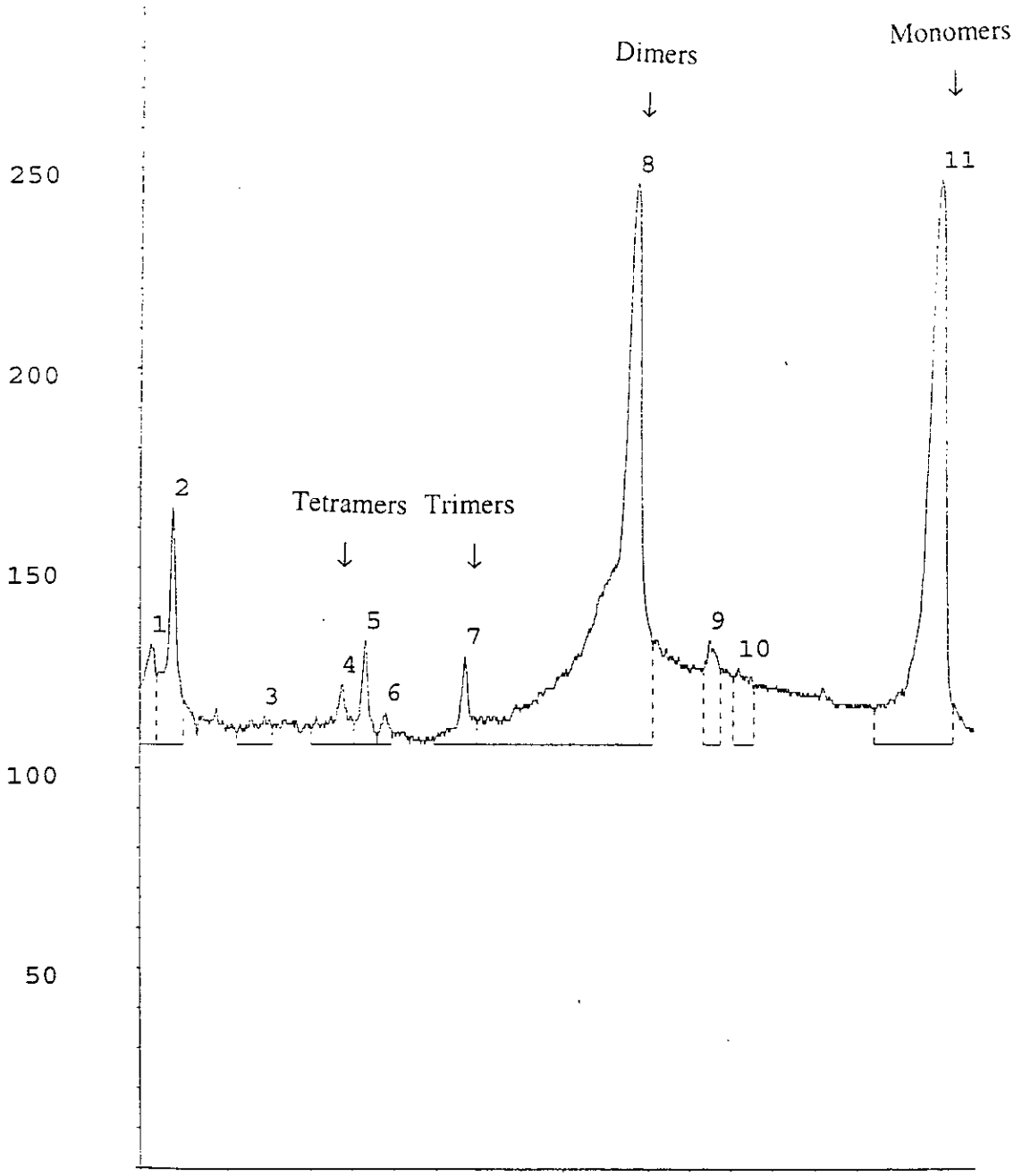


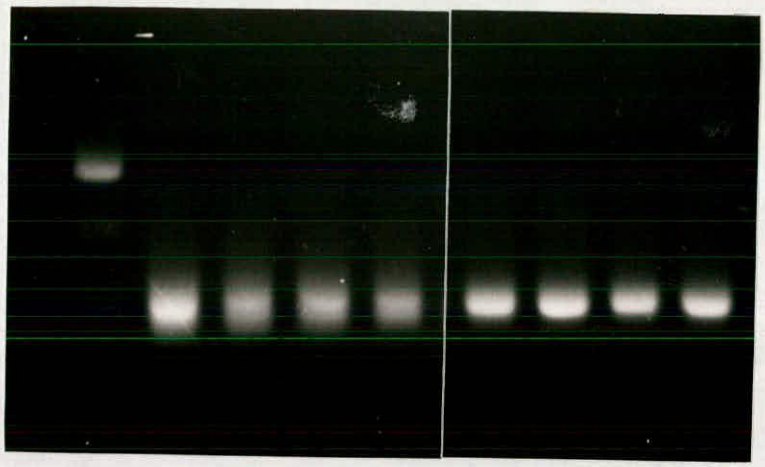
Figure 21. Distribution of plasmid multimers in pUC18



pMS7, the remaining colonies were all blue. Transformants were obtained from pAC2 that appeared to sector. The colonies were purified by streaking to give single colonies, liquid cultures produced and plasmid DNA extractions performed. The extracted DNA was electrophoresed on a 1% agarose gel overnight. The single colony obtained for pMS7 seemed to contain plasmid of the correct size to be a monomer and still contain the palindrome (figure 22). The DNA from the pAC2 colonies were also monomers but the DNA formed a diffuse band running at a point between pMS7 that contained the palindrome and monomers of pUC18 .

An *sbcCD* strain (DL733) was then transformed with the artificially created monomeric DNA to try and determine whether the multimerisation observed in this host was a consequence of the host genotype or the presence of the palindrome. Transformation with pMS7 resulted in a higher frequency of blue colonies indicating that the palindrome was less stable when the plasmid was in monomeric form. Transformations were also performed with pAC2 and pUC18 monomers. DNA was extracted and the uncut plasmid electrophoresed on a 1% agarose gel. The DNA from artificially created pMS7 monomers was introduced into an *sbcCD* host ran as a monomer but bands corresponding to dimers and trimers were also visible (figure 23). pAC2 and pUC18 appeared to be of the same size indicating that the production of artificially created monomers had resulted in the loss of the palindrome from pAC2. Both plasmids were present as monomers along with multimeric forms.

Figure 22. Artificially created plasmid monomers in a *recA sbcCD* strain.



1 2 3 4 5 6 7 8 9

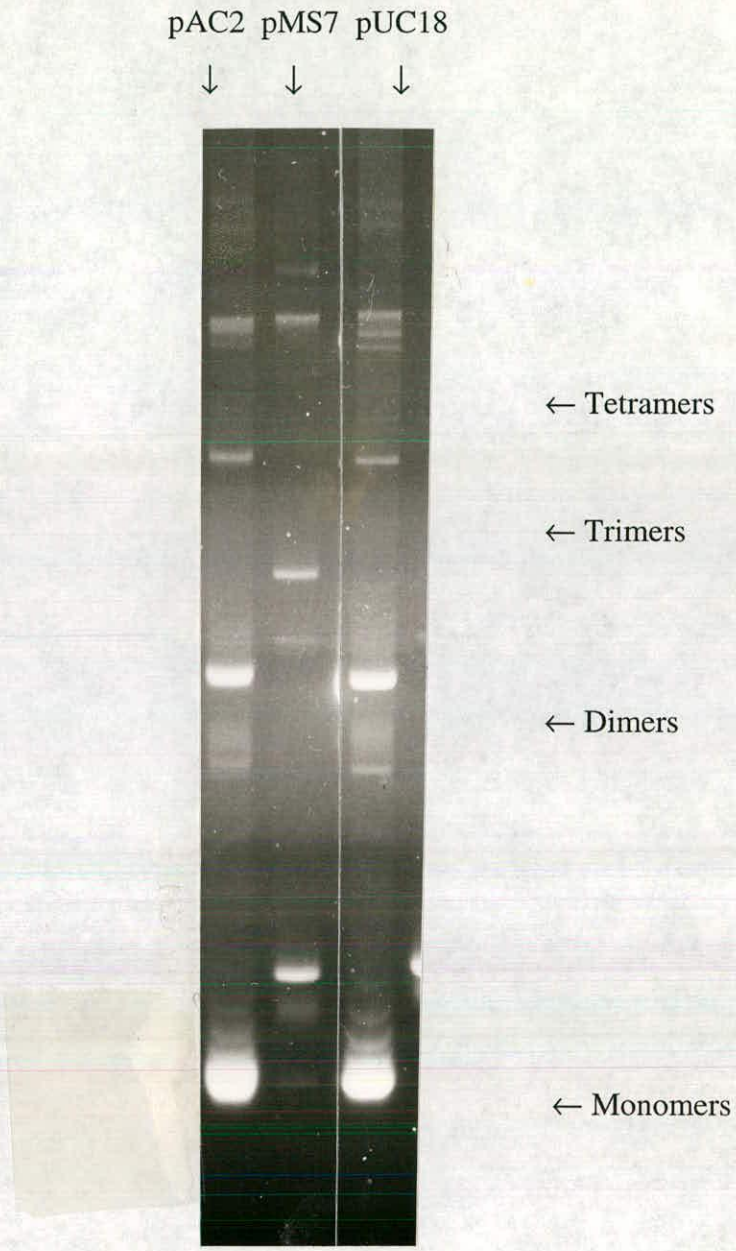
Lanes

1. pMS7

2-5. pAC2

6-7. pUC18

Figure 23. Plasmid monomers after re-introduction into an *sbcCD* background.



The effect of host genotype on the deletion frequency on the shorter 109 bp palindrome was investigated by transforming isogenic *E. coli* strains with pDLJ2. When a wild-type *E. coli* strain (DL324) was transformed colonies were obtained since the 109 bp palindrome was below the size limit for inviability and so could be tolerated. The colonies were convex in shape and were covered with light and dark blue flecks corresponding to cells containing plasmids which had deleted the palindrome. When an *sbcC* or *sbcCD* strain (DL494 or DL733 respectively) was transformed the deletion frequency was lower than that observed in the wild-type. Many of the colonies appear 'sick' with concave centres and produced large white sectors when stored on a plate for a seven days at room temperature. Approximately 25% of the colonies were white. Plasmid DNA from these white colonies still carried the palindrome and the colonies remained white when the DNA was re-introduced into an *sbcCD* background.

The deletion frequency was higher than in a *recA* background than in the wild-type strain, as determined by colony morphology. This could have been due to the lack of recombinational repair of stalled replication forks in this background. In morphology, the colonies were shiny and convex, similar to the wild-type. Introduction of the *recA* mutation, by P1 transduction, into an *sbcC* or *sbcCD* background abolished the poor colony morphology and the high frequency of white colonies seen in *sbcCD* strains. The level of palindrome deletion appeared similar to that of a wild-type background. The poor colony morphology observed in an *sbcCD*

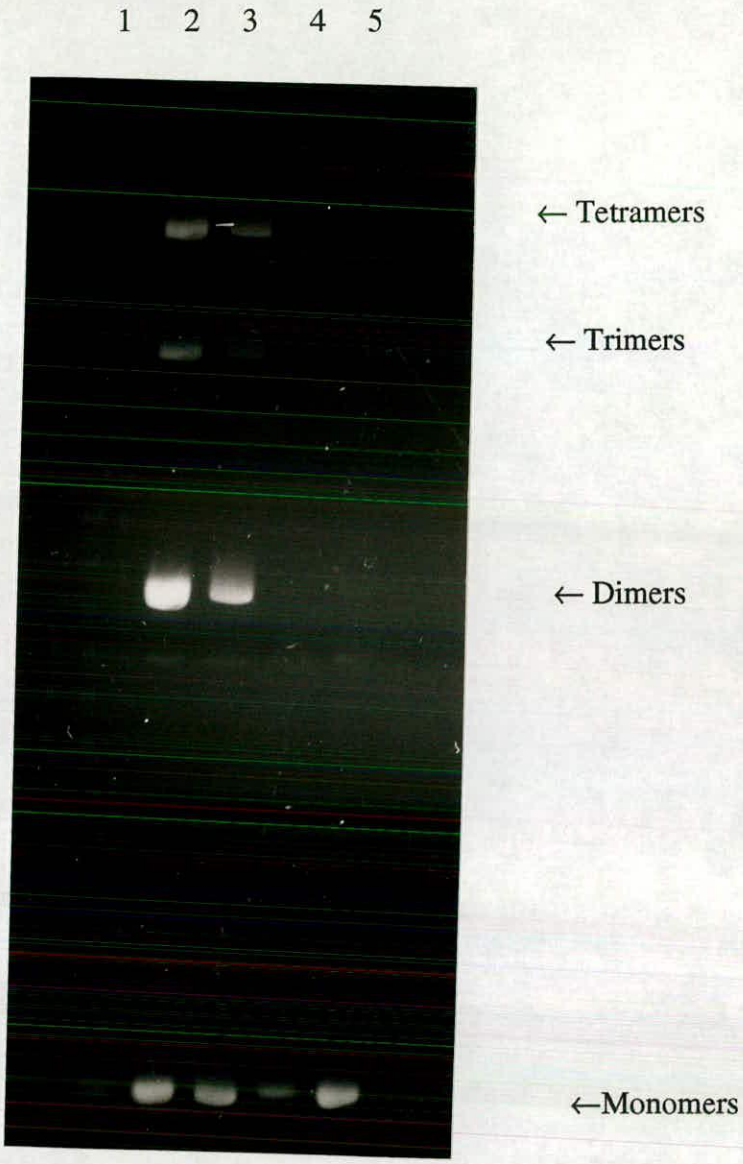
background could have been a result of SOS induction. RecA initiates the SOS response by binding to single-stranded DNA (Kowaczykowski *et al.*, 1994). It may be that the presence of single stranded regions of DNA at stalled replication forks that activates the SOS pathway and the introduction of a *recA* mutation into the *sbcCD* background would prevent this induction. The introduction of a *sfiA* mutation (which prevents SOS induced inhibition of cell division) into the *sbcCD* background by P1 transduction should mimic the effect seen with the *recA* mutation if the poor colony morphology is a result of inhibition of cell division caused by SOS induction. However the *sfiA* mutation did not alleviate the poor colony morphology.

DNA was purified from each of the strains containing the plasmid pDLJ2 and electrophoresed uncut on agarose gels to separate the different topoisomeric forms of the plasmid (figure 24). When DNA from the strains containing pDLJ2 were examined it was apparent that the yield of DNA was low in the wild-type background. This was consistent with the difficulties associated with replicating palindromic DNA in this background. The DNA that was present appeared to be in a monomeric form. In an *sbcC* or *sbcCD* background the yields of DNA were higher possibly indicating that there was a residual level of inviability associated with the palindrome in the wild-type background even though the palindrome was below the size limit for complete inviability. Most of the plasmid in the *sbcC* and *sbcCD* backgrounds were present in multimeric forms, in particular dimers, trimers and tetramers. In a *recA* background the plasmid DNA was present as monomers and a

faint band running ahead of the monomers corresponded to the deleted form of the plasmid. This was consistent with the elevated levels of deletion in this background, as determined by examination of the colony morphology. The absence of multimeric forms in this background was to be expected since recombination between plasmid molecules was inhibited because of the lack of a functional RecA protein. In a *recA sbcCD* host the DNA was again in a momomeric form but no band corresponding to deletion products was visible. This was consistent with the lower deletion frequency deduced from the colony appearance.

Although no palindrome free pUC119 was available as a control, comparison can be made between the behaviour of pDLJ2 and pUC18. In a wild-type strain pDLJ2 was present as monomers and the DNA content was low. In contrast pUC18 was present as monomers but also higher multimeric forms (figure 25). In an *sbcCD* background both plasmids were present as multimers (figure 25). In a *recA* or *recA sbcCD* strain the DNA for both plasmids was present predominantly as supercoiled monomers although some faint bands were present corresponding to relaxed monomers and possibly some dimers (figure 26).

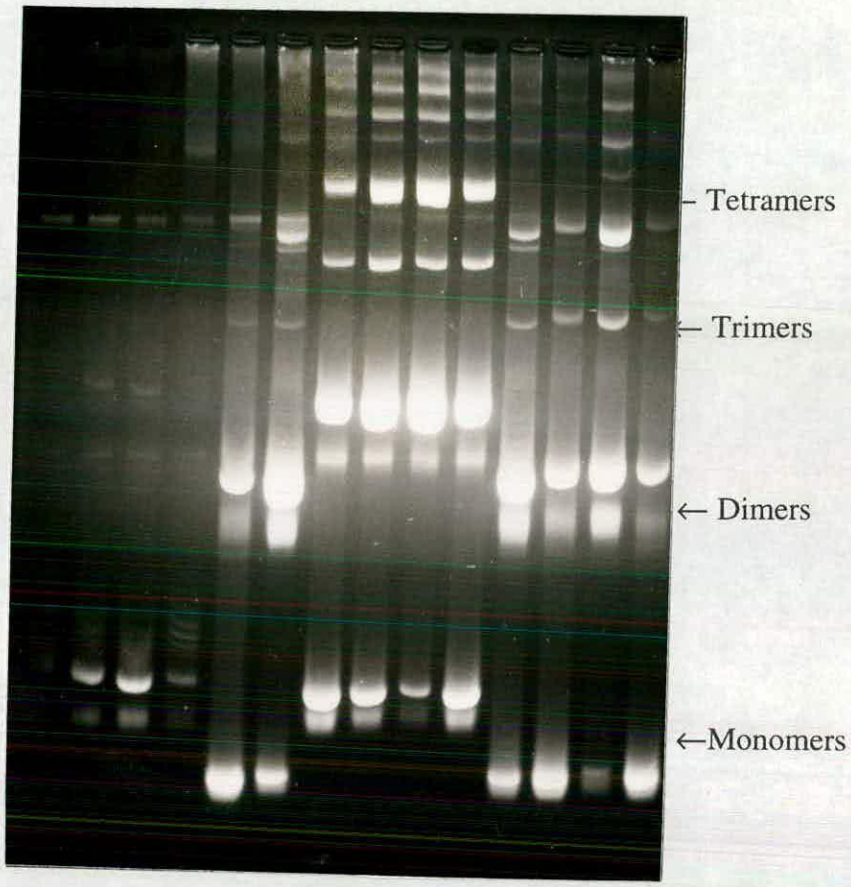
Figure 24. pDLJ2 DNA extracted from different *E. coli* backgrounds.



Key

1. DL324 (wild-type)
2. DL494 (*sbcC*$^-$)
3. DL733 (Δ *sbcCD*)
4. DL887 (*recA*)
5. DL888 (*recA sbcCD*)

Figure 25. pDLJ2 and pUC18 DNA purified from wild-type and *sbcCD* strains



Key

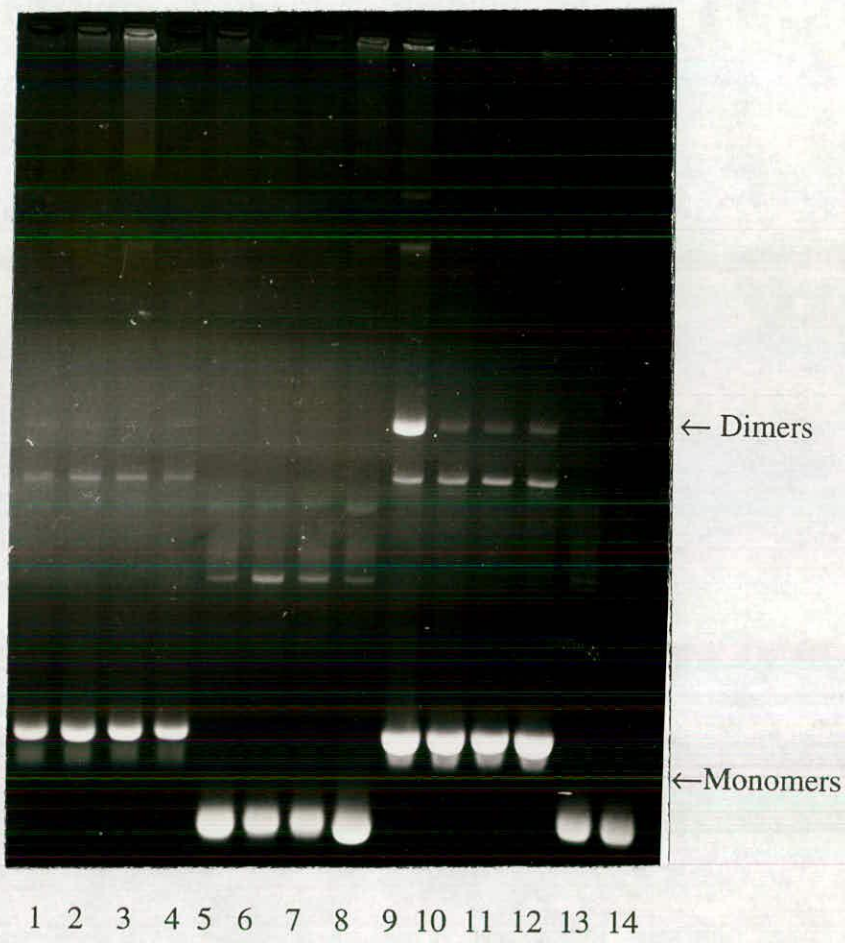
1-4. pDLJ2 in JM83 (wild-type)

5-6. pUC18 in JM83 (wild-type)

7-10. pDLJ2 in DL733 (*sbcCD*)

11-14. pUC18 in DL733 (*sbcCD*)

Figure 26. pDLJ2 and pUC18 DNA purified from *recA* and *recA sbcCD* strains



Key

1-4. pDLJ2 in DL887 (*recA*)

5-9. pUC18 in DL887 (*recA*)

9-12. pDLJ2 in DL888 (*recA sbcCD*)

13-14. pUC18 in DL888 (*recA sbcCD*)

Discussion.

Factors affecting the deletion of long palindromic sequences from plasmid vectors were investigated. A 571 bp near perfect palindrome was cloned into the high copy number plasmid pUC18 to give the plasmid designated pAC2 (Chalker, 1990). The palindrome apparently deleted at a frequency of approximately 50%. This high level of the deletion may be due in some part to the high copy number of the plasmid and the tendency of palindrome free derivatives of the plasmid to have a replication advantage over plasmids bearing the palindrome. The central 109 bp palindrome present within the 571 bp palindrome was previously subcloned into pUC18 and shown to be stable, but was unstable in M13 mp18 (Leach *et al.*, 1987). This difference in stability may reflect some aspect of the local sequence context of the palindrome, a long range sequence effect or the fact the M13 is a single stranded filamentous phage. The single stranded nature of the phage could increase the chances of the palindrome adopting a hairpin structure which would promote replication slippage during the double stranded replication cycle of the phage. It may be that there is a more general problem associated with the cloning of palindromes in M13 and not just with the 109 bp palindrome.

The central 109 bp palindrome could be stably cloned into the plasmids pUC118 and pUC119 but showed a degree of instability in pUC119 (D.Leach, unpublished). Using these plasmids it was demonstrated that their was a bias for

deletion to occur on the lagging strand of the replication fork. The preferred position of the direct repeats used in the deletion occurred on the lagging strand of pUC119 but on the leading strand of pUC118. This observation of a bias in the deletion supports the work performed previously by Trinh and Sinden (1991) which used a different system than the one used in this study. However, it is unclear why the system used by Weston-Hafer and Berg (1991) failed to find a bias for either the leading or lagging strand. Both groups used a system which employed the cloning of palindromes into lower copy number plasmids (pBR325 and pBR322 respectively). It may be that the discrepancy reflects some feature of the insertion sites used.

An investigation was made into a series of plasmid derivatives of pAC2 which exhibited the stable maintenance of a long palindromic sequence. Despite employing a number of different methods no differences at the sequence level could be detected. However, it was shown that the two plasmids pMS5 and pMS7 were less negatively supercoiled than the parental plasmid pAC2 and the palindrome free plasmid pUC18. It is unclear whether the differences in supercoiling levels promotes stabilisation of the palindrome, or if the more relaxed nature of the plasmids is a consequence of the increased stability.

Subsequent work involving the plasmids used in this study has revealed new information about the structure of plasmid pAC2 (D.Pinder and D.Leach, personal communication). It was originally assumed that pAC2 existed as a monomer or

homodimer. Deletion of one copy of the palindrome could occur from the dimer resulting in a heterodimer, which might stabilise the remaining palindrome. However, digestion of the plasmids pAC2, pMS5, pMS7 and pUC18 with the enzyme *Pvu*II revealed the presence of a larger than expected fragment for pAC2 in the region between base numbers 306-628 from the half of the plasmid dimer which was assumed to have deleted the palindrome. This fragment was not detected in the initial restriction fragment analysis of the plasmids because the four base cutter enzymes used placed this region of the plasmid on restriction fragments which were small in size and poorly resolved from fragments of similar size on the agarose gel (table 6). Sequence information for this indicates the presence of a short directly repeated sequence between the *Eco*RI and *Bam*HI sites in the multiple cloning site. Re-circularistion of the plasmid bearing this repeat following introduction into an *sbcCD* background gives rise to sectored colonies, possibly indicating an unstable insertion of DNA. The presence of this region within the normal heterodimer of pAC2 would also explain why pAC2 gives rise to sectored colonies. The deletion of one copy of the palindrome from a dimer should give rise to a blue colony rather than the sectored colonies actually observed unless the expansion/contraction of this repeated sequence switched the plasmid between lac^+ and lac^- status.

The presence of a repeated sequence within the multiple cloning site in one half of the dimer could also explain the diffuse band pattern observed when artificially created monomers of pAC2 were electrophoresed on an agarose gel (see

figure 22). When the monomers were cut with the enzyme *Hind*III two bands were seen for the monomer of pAC2 compared with the single band observed for pUC18 (figure 27). This pattern of double bands was also seen when the uncut monomeric DNA was electrophoresed on a gel containing a low concentration of chloroquine (figure 28). The formation of this short, unstable repeated sequence within pAC2 might act to stabilise the remaining copy of the palindrome however this region has not been detected in the plasmid pMS5 and pMS7 which exhibit greater palindrome stability so the nature of the stability in these plasmids is still unknown.

An investigation was made into the effect of different host genotypes on the propagation of plasmids bearing long palindromes. Plasmids pAC2, pMS5 and pMS7 bearing a 571 bp palindrome could be recovered from an *sbcCD* but the majority of the DNA was in a multimeric form. When the DNA was introduced into a *recA sbcCD* strain in a multimeric form (predominantly dimeric) it remained in this form. It was therefore decided to look at the fate of artificially created monomers in both of these backgrounds. Plasmid pAC2 could not be stably propagated and deletion products accumulated. The plasmid derivative pMS7 could be isolated as a monomer but only with difficulty and the DNA isolated may have accumulated mutations to allow it to exist as a monomer.

The plasmid pDLJ2 bearing a 109 bp palindrome could be propagated in all the backgrounds tested. In a wild-type strain the yields of DNA were low and the

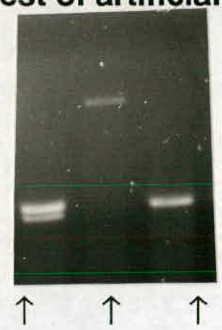
plasmid existed purely in a monomeric form. In an *sbcCD* strain the plasmid was present as a monomer but the majority of the DNA was in a multimeric form. In a *recA* and *recA sbcCD* strain the plasmid was present predominantly in a monomeric form. It would appear that the presence of palindromic sequences in a plasmid leads to the formation of plasmid multimers in an *sbcCD* background

Table 6. Restriction fragments produced by 4 base cutters.

*Alu*I (AG'CT)	*Hha*I (GCG'C)	*Hae*III (GG'CC)
679	902	828
521	393	587
257	337	458
245	332	434
226	**331**	298
165	270	267
165	174	**257**
113	130	137
113	109	102
100	103	80
95	100	37
93	100	18
90	93	11
64	67	
64	65	
64	30	
63	28	
54		
46		
45		
25		
19		

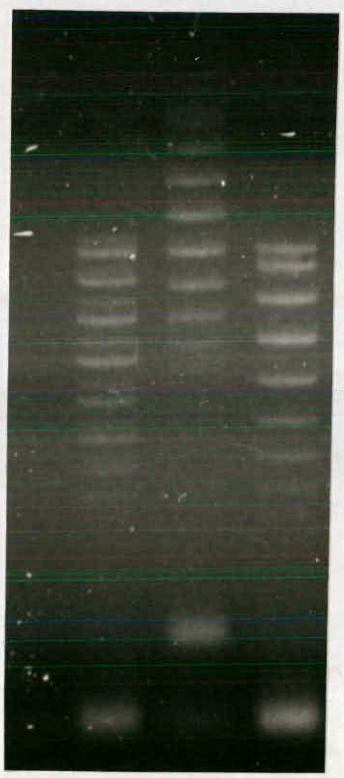
Numbers shown in bold correspond to those restriction fragments which include the short repeated sequence.

Figure 27. *Hind*III digest of artificially created monomers.



pAC2 pMS7 pUC18

Figure 28. Chloroquine gel of artificially created monomers.



pAC2 pMS7 pUC18

CHAPTER 4

INFLUENCE OF *SBCCD* MUTATION ON DNA REPLICATION

Introduction

Rapidly growing wild-type strains of *E. coli* can duplicate in a time much shorter than the time required for chromosomal replication (Donachie, 1992). This results in multiple replication bubbles in the chromosome. Under these conditions it is essential that the integrity of the replication forks is maintained. The DNA strands that serve as templates for replication must be free of irregularities and interruptions. Occasionally, replication forks do collapse and are subject to recombinational repair. It has been proposed that in *E. coli,* disintegrated replication forks are repaired by the combined action of RecA and RecBCD proteins (Kuzminov, 1995).

The proposed model involves a theta replicating chromosome which, upon reaching a single-stranded break, collapses to produce a sigma structure. The free double-strand end is degraded by the RecBCD enzyme until its nuclease activity is inactivated by a Chi site present in the correct orientation in the DNA strand (Stahl, 1990). At this point the RecBCD enzyme loses its nuclease activity, possibly because of dissociation of the RecD subunit, and proceeds as a helicase unwinding the DNA duplex. RecA protein then forms a filament coating the single-strand end and promotes invasion of the intact homologous duplex, leading to the formation of a Holliday-junction. Resolution of the junction by RuvC restores the replication fork (Connolly *et al.*, 1991).

If both replication forks of a replication bubble collapse because of nicks in the same strand it could result in the release of a whole replication fork as a linear molecule. This molecule would be susceptible to RecBCD degradation. The RecBCD-dependent degradation of the chromosomes would be especially evident in *recA* strains which cannot repair collapsed replication forks and so would lead to no net increase in the chromosome number from that replication bubble.

The degradation of chromosomes by RecBCD nuclease is thought to have been detected in a previous study utilising flow cytometry technology (Skarstad and Boye, 1993). When rapidly growing cultures of wild-type *E. coli* are treated with rifampicin, which allows completion of ongoing rounds of replication without the initiation of new rounds, the cells are seen to contain two, four or eight fully replicated chromosomes (Skarstad and Boye, 1986). *RecA* strains, which are recombination deficient, exhibit an asynchrony phenotype in which cells contain three, five, six or seven fully replicated chromosomes. The inter-initation time in these strains is normal, indicating that initiation of replication is essentially normal and occurs at a specific time in the cell cycle (Skarstad and Boye, 1988).

A model to explain this phenomenon was suggested in which there was selective degradation of individual chromosomes, before and during rifampicin treatment, by RecBCD nuclease (Skarstad and Boye, 1993). The replication phenotypes of a wild-type, *recA*, *recD* and a *recA recD* strain were investigated by

treatment with rifampicin followed by flow cytometry. The wild-type and *recD* strains contained two and four chromosomes and the *recA* strain contained one, two, three and four fully replicated chromosomes. Cells of the double mutant contained mainly two and four chromosomes. It was suggested that the exonuclease activity of the RecBCD enzyme was responsible for the chromosome degradation seen in *recA* strains. DNA degradation was measured directly in these strains after pre-labeling of cultures with [^{3}H]thymidine and comparison of the amounts of acid-insoluble label in samples withdrawn at varying times after the addition of rifampicin. It was shown that there was little degradation in the wild-type, *recD* and *recA recD* strains but approximately 40% degradation in the *recA* strain during the period of incubation with rifampicin. Replication forks may also be stalled by the presence of secondary structures within the DNA to replicated. As discussed in Chapter 1 it has been proposed that an endonuclease could recognise some feature of this structure and remove it to produce a broken replication fork with a free end. Exonuclease digestion of part or all of the chromosome arm would remove the secondary structure. It has been demonstrated that the *SbcCD* gene products of *E. coli* possess an ATP dependent double-strand exonuclease activity. This nuclease activity may be responsible for the removal of secondary structures from stalled replication forks (Leach, 1994). Recombination between the partially digested arm and a homologous sequence impedes unnecessary degradation of the chromosome.

It was decided to investigate whether the degradation of stalled replication forks by SbcCD nuclease does take place in *E. coli*, using flow cytometry. Also, whether the asynchrony phenotype associated with *rec*A strains is alleviated in a *recA sbcCD* background. The following work was performed with the kind help of Dr Eric Boye and Dr Kirsten Skarstad.

Results

The cultures of interest were grown as previously described with rifampicin added to the culture medium to prevent initiation of fresh rounds of replication. Cephalexin was also added to prevent cell division and so increase the total number of chromosomes present in individual cells. Using flow cytometry a comparison was made of both exponentially growing (pre-rifampicin treated) and rifampicin treated cultures. The results are presented in histogram form with comparisons of the cell number and chromosome number of cultures shown.

In the strain DL473 (which has a wild-type recombination genotype) a wide distribution of chromosome numbers was visible in the exponentially growing culture with cells containing incompletely replicated chromosomes (figure 29a). Upon rifampicin treatment ongoing rounds of replication were completed to produce cells with mainly four or eight fully replicated chromosome equivalents (figure 29b). The isogenic *recA* derivative, DL519, gave a histogram in which the cells containing four and eight chromosomes had undergone degradation reducing the number of chromosome equivalents to three, two and one, or seven, six and five respectively (figure 29d). The untreated culture of this strain contained a proportion of cells with no DNA (figure 29c) represented by a peak to the far left of the histogram.

In the *recA recD* double mutant (DL520) the peaks visible in the *recA* strain were much less distinct with a major peak present at four chromosome equivalents (figure 29f). It could be argued that rather than there being less degradation there was actually more and this led to cells with partial chromosomes resulting in a graph which looked more like that of an exponentially growing culture.

A *recA sbcC* double mutant (DL521) appeared identical to the *recA* strain in so far as there were peaks present at one, two, three, four, five, six, seven and eight chromosome equivalents (figure 29h). This indicates that the introduction of an *sbcC* mutation does not alleviate the asynchronous phenotype seen in a *recA* background. A similar result was obtained in the *E. coli* backgrounds DL888 and NM772 indicating that the result was not background specific (data not shown). When the triple mutant *recA recD sbcC* (DL522) was examined no distinct peaks were evident (figure 29j). Instead there was a broad distribution of chromosome numbers with a proportion of cells containing more than eight chromosome equivalents and less cells with one or two chromosomes. In the corresponding exponentially growing culture (figure 29i), the cells appeared to contain very little DNA, approximately 25% of that present in the wild-type strain. However, the rifampicin treated *recA recD sbcC* cells had a normal DNA content. This phenomenon was also seen in another *recA recD sbcC* background, DL847(figure 29l). When exponentially growing cultures of both DL522 and DL847 were stained using the DNA specific dye DAPI, and examined using fluorescence microscopy, the DNA content appeared normal (data not shown). It may be that this apparent lack of DNA in the untreated cells is an

artefact due to the DNA not being properly stained with mithramycin, the DNA specific dye used in flow cytometry. However it is unclear why the triple mutant should behave differently from any of the other strains with respect to staining.

Figure 29. DNA histograms of cell number versus number of chromosome equivalents.

29a. Exponentially growing culture of DL473 (wild-type)

29b. Rifampicin treated culture of DL473

29c. Exponentially growing culture of DL519 (*recA*)

29d. Rifampicin treated culture of DL519

29e. Exponentially growing culture of DL520 (*recA recD*)

29f. Rifampicin treated culture of DL520

29g. Exponentially growing culture of DL521 (*recA sbcC*)

29h. Rifampicin treated culture of DL521

29i. Exponentially growing culture of DL522 (*recA recD sbcC*)

29j. Rifampicin treated culture of DL522

29k. Exponentially growing culture of DL847 (*recA recD sbcC*)

29l. Rifampicin treated culture of DL847

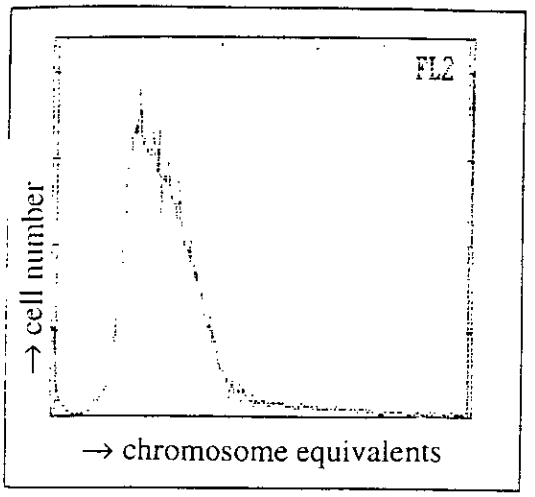


Figure 29a.

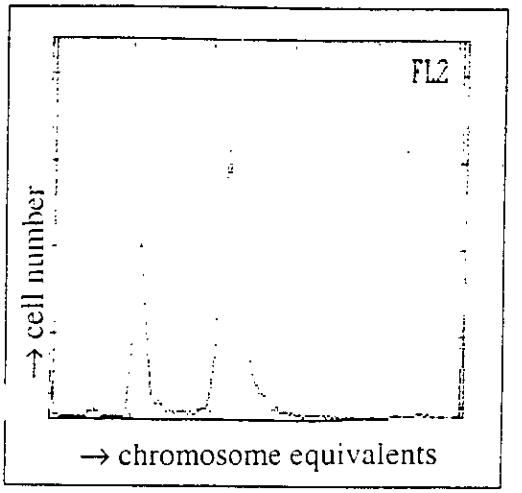


Figure 29b.

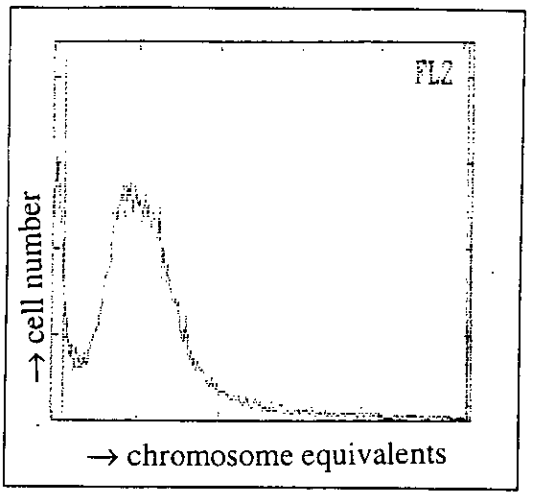


Figure 29c.

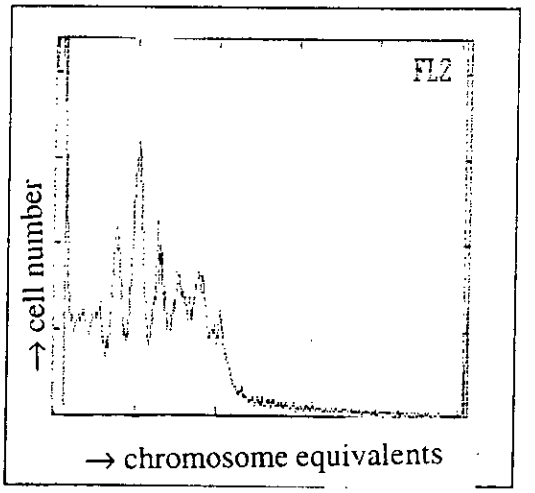


Figure 29d.

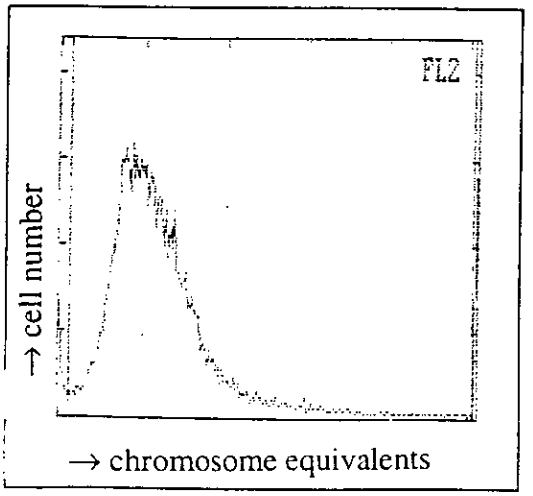


Figure 29e.

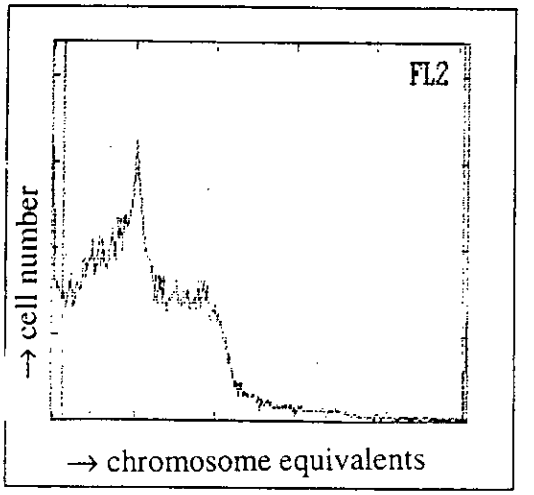


Figure 29f.

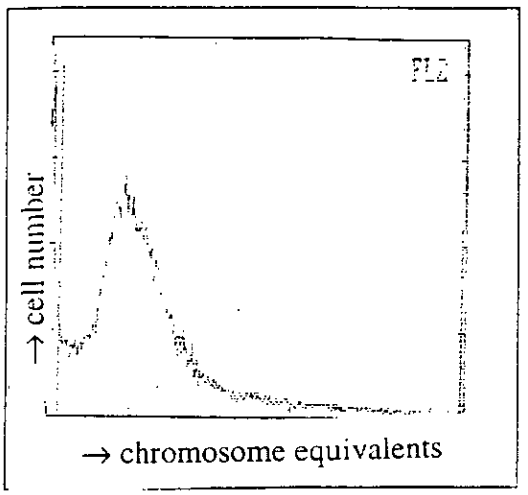


Figure 29g.

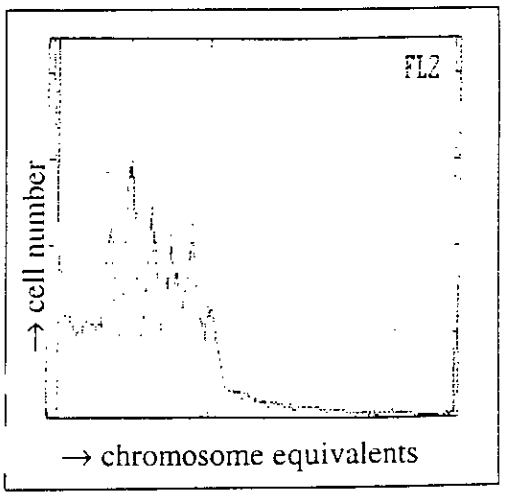


Figure 29h.

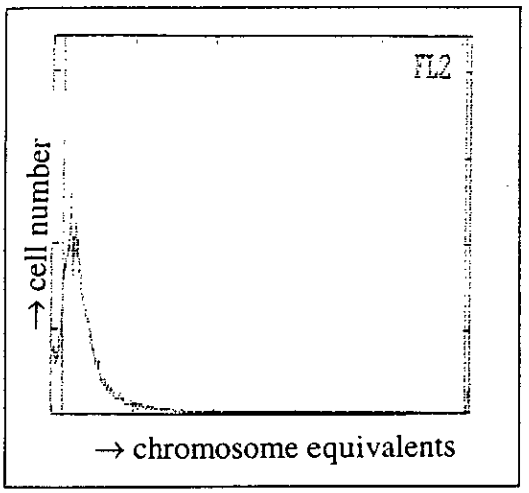


Figure 29i.

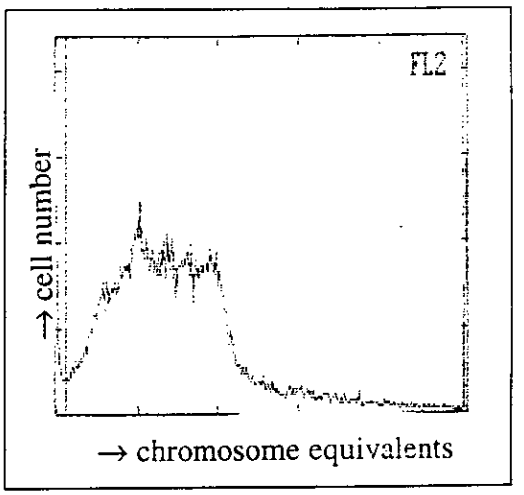


Figure 29j.

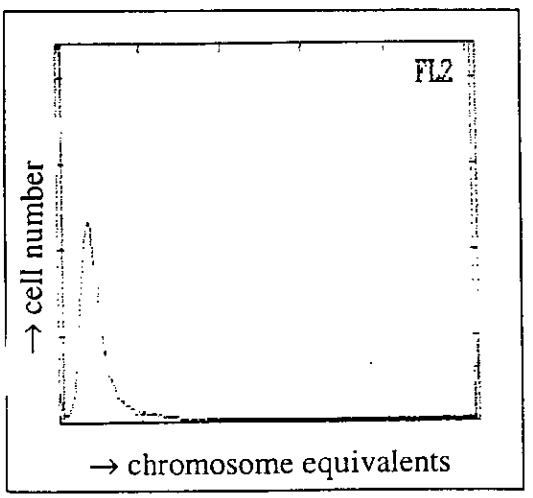


Figure 29k.

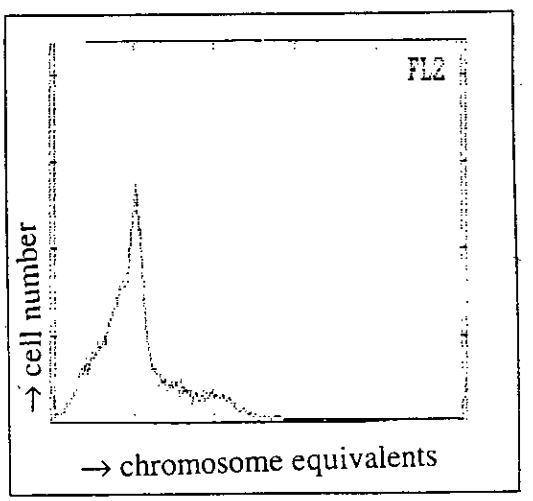


Figure 29l.

Discussion

The data presented here suggest that the asynchrony phenotype associated with a *recA* strain was not alleviated in a *sbcC* background. However, the SbcCD proteins are thought to act on secondary structure present within the replicating DNA (Leach, 1994) and since long palindromic sequences which may adopt these structures are absent from the *E. coli* genome, chromosome degradation by SbcCD may be a rare event undetectable by this method. The introduction of a long palindromic sequence onto the chromosome may facilitate the study of this type of secondary structure repair mechanism. It may also be that inactivation of the exonuclease activity of the RecBCD enzyme by the use of a *recD* mutation may allow degradation of the chromosome by other nucleases present within the cell, many of which may be more potent and act before SbcCD. It could also be postulated that there may be more than one gene product and more than one mechanism involved in the breakdown of replication forks and so the effect of SbcCD will be masked. One feature of RecBCD recombination that was not discussed in the earlier work performed by Skarstad and Boye (1993) is the involvement of Chi sequences in arresting DNA degradation by RecBCD. If the chromosomal degradation associated with the asynchrony phenotype is a result of RecBCD nuclease activity it might be expected that the degradation may be arrested by Chi site activity before a whole chromosome arm is lost. Accumulating data on *E.coli* sequences show that the Chi sequence (5'-GCTGGTGG-3') is overrepresented in the genome, but that it is eight to nine times more abundant in the orientation leading towards the origin of

replication (Burland et al., 1993). It might be expected that the action of Chi sites on RecBCD would take place before the loss of an entire chromosome unless there is degradation attributable to another nuclease.

CHAPTER 5

ARAB-LACZ FUSION FORMATION IN *E. COLI*

Introduction

The *ara-lac* fusion system has been one of the best documented examples of directed mutation, in which mutations do not arise during the normal growth of the culture but do appear when the culture is plated on selective medium (for reviews see Foster, 1992 and 1993). Shapiro (1984) used the strain MCS2 to analyse the formation of *araB-lacZ* coding sequence fusions. The strain has a *Mu* prophage DNA inserted between the *araB* and *lacZ* genes. Excision of the prophage fuses the two genes and if the fusion occurs in-frame it will place the *lacZ* gene under the control of the *araB* promoter. Plating of the strain on medium containing lactose as a growth substrate and arabinose as an inducer gives rise to the formation of colonies after an initial lag period. After that time colonies rapidly appear. Colonies are never detected in the first two days after plating indicating that fusion colonies are not present in the initial culture. The fusion point can occur in-frame in a 26 bp segment of the *lacZ* gene between nucleotides 49 and 75. For fusion to occur there is a requirement for the Mu A transposition function (Shapiro and Leach, 1990) and the excision involves a complex DNA rearrangement (Shapiro and Leach, 1990). Certain host encoded functions are also required, including IHF (host integration factor), HU and ClpPX protease (Shapiro and Leach, 1990; Shapiro, 1993).

Given the role of complex DNA rearrangements in excision of the prophage and the possible involvement of secondary structure, an investigation was made into

the effect of *sbcCD* mutation alone and in conjunction with *mutS* , on *araB-lacZ* fusion formation. The original fusion strains used in this work were kindly provided by Dr. Genevieve Maenhault-Michel.

Results

Two clones, MCS2a2 and MCS2a3 (derived from the original MCS2 strain), were patched onto minimal glucose medium and incubated overnight. The patch of growth was scraped off the plate with a sterile toothpick and resuspended in 1ml of PA salts. 100 µl aliquots of a 10^{-2} and 10^{-4} dilution of the cell suspension were plated onto medium containing arabinose and lactose and the plates incubated overnight at 30°C. No fusion colonies were detected until approximately 9 days after plating indicating that no fusions were present in the original culture. Over a period of days the number of fusion colonies formed increased (figures 30, 31, 32 and 33). No difference was observed between the two dilutions demonstrating that the number of fusions formed was independent of the total number of cells plated, as observed by other workers (Shapiro, 1984; Maenhault-Michel and Shapiro, 1994).

To investigate the involvement of chromosomal mutations on fusion formation, derivatives of MCS2a2 were constructed which carried mutations in *sbcCD* (DL943), *mutS* (DL944) and *sbcCD mutS* (DL945). The strains were plated out as previously described and the formation of fusion colonies noted. No difference in the kinetics of fusion appearance was observed for any of the strains bearing chromosomal mutations as compared with the wild-type strain (figures 34, 35, 36 and 37).

Figure 30. Ara-Lac fusion colony formation in a wild type host MCS2a2 (plated at 10-2 dilution)

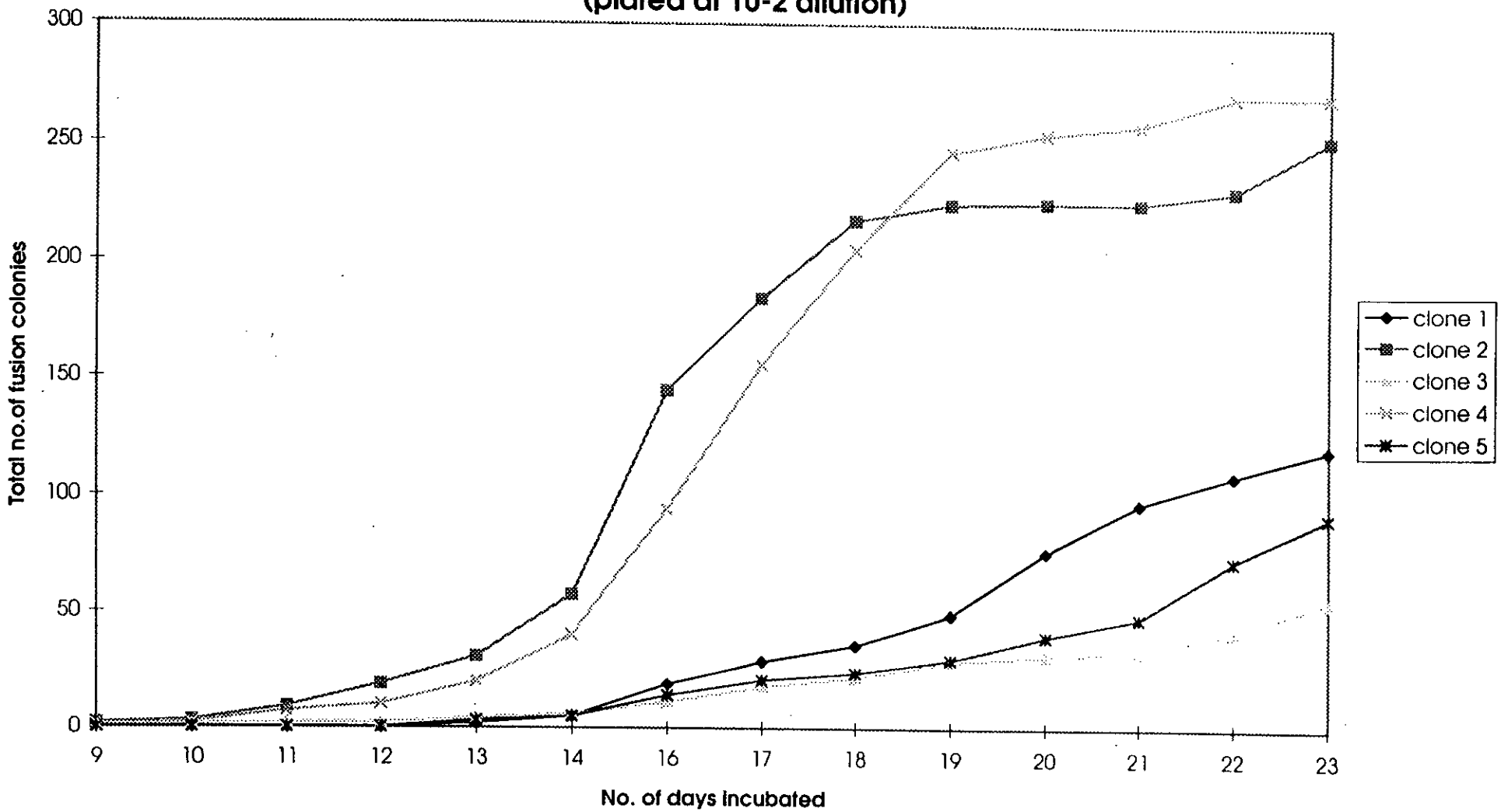


Figure 31. Ara-Lac Fusion Colony Formation in wild-type host MCS2a2 (plated at 10-4 dilution)

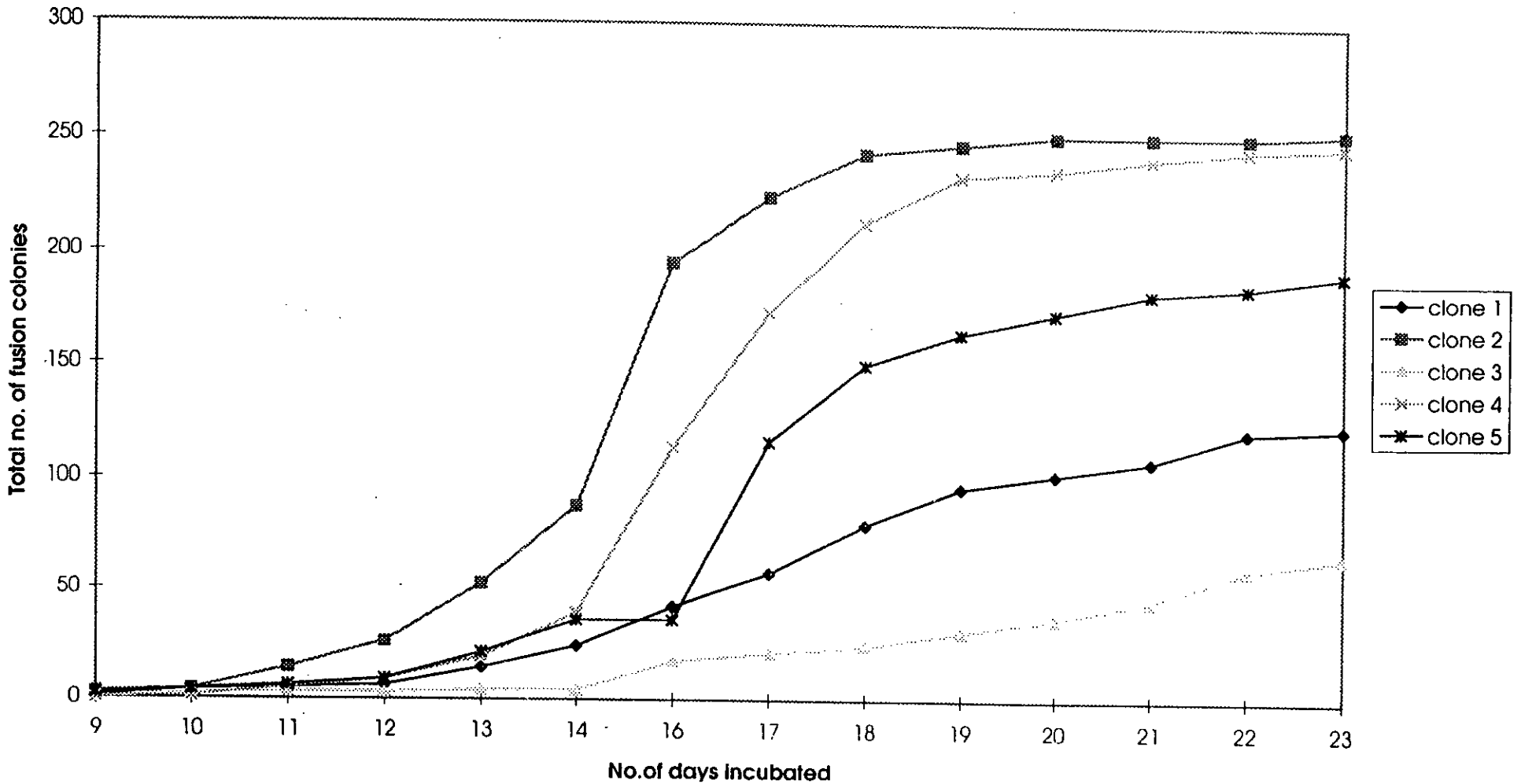


Figure 32 .Ara-Lac fusion colony formation in a wild type host host MCS2a3 (plated at 10-2 dilution)

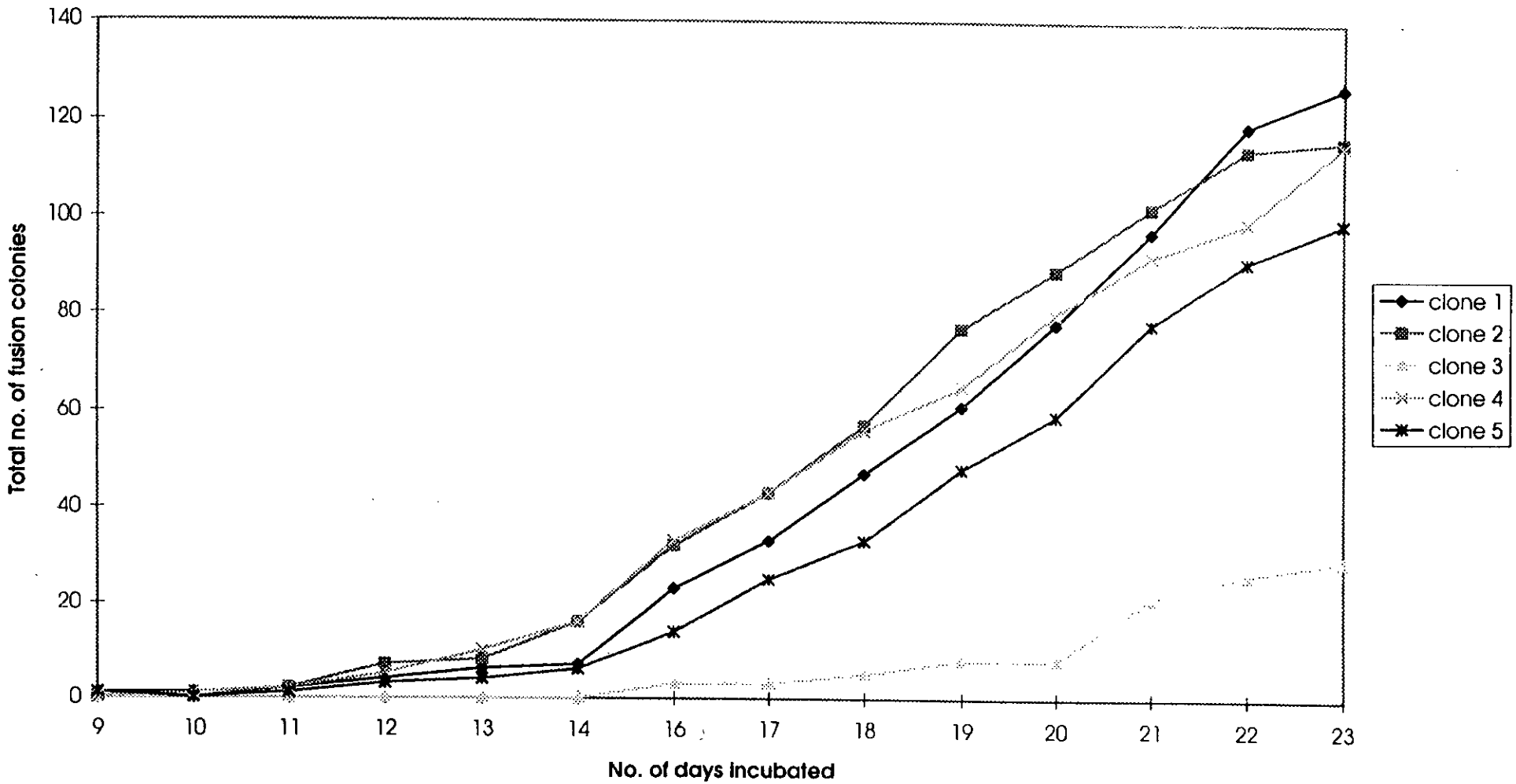


Figure 33. Ara-Lac fusion colony formation in a wild-type host MCS2a3 (plated at 10-4 dilution)

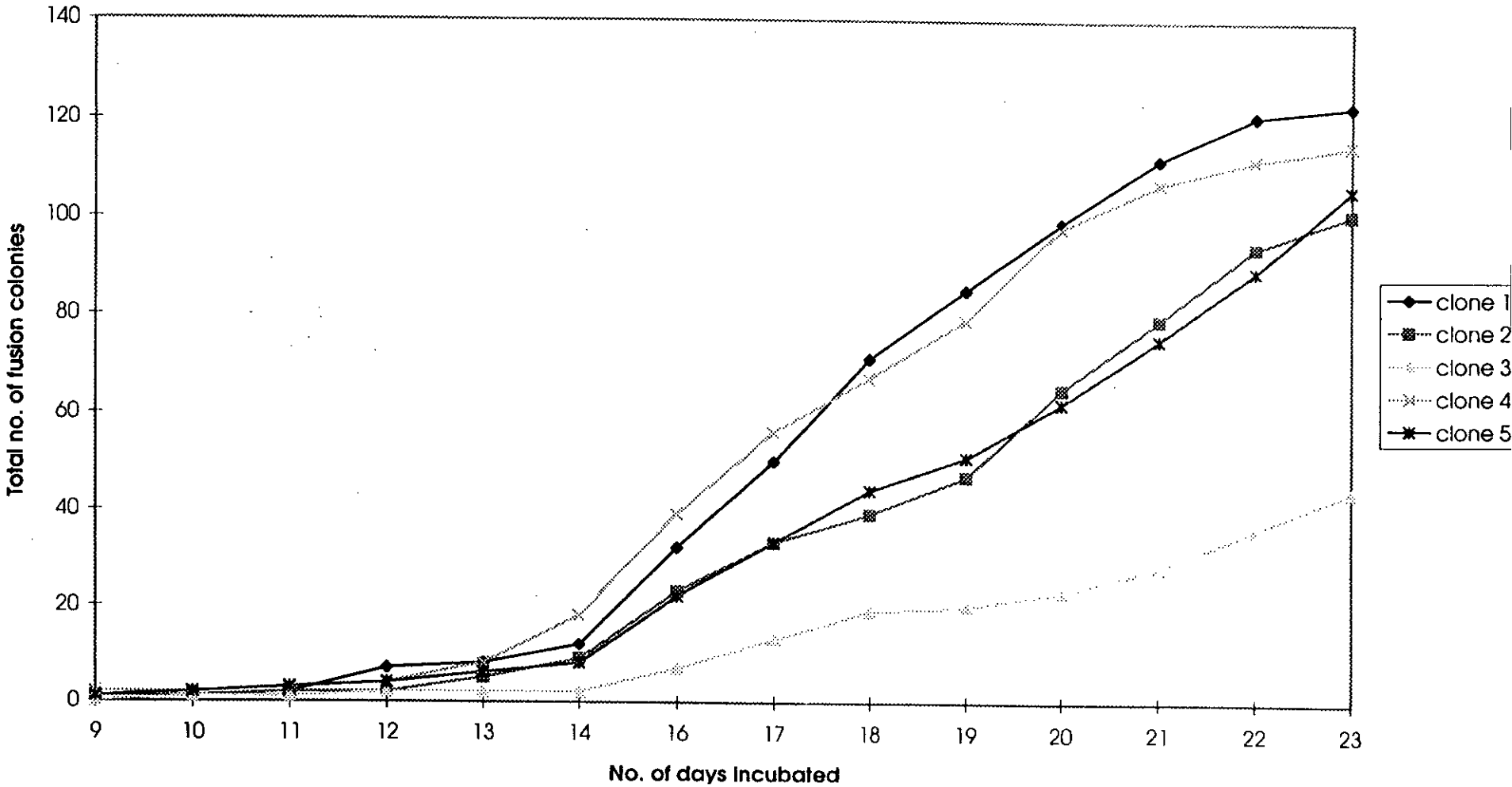


Figure 34. Ara-Lac fusion colony formation in MCS2a2 wild-type

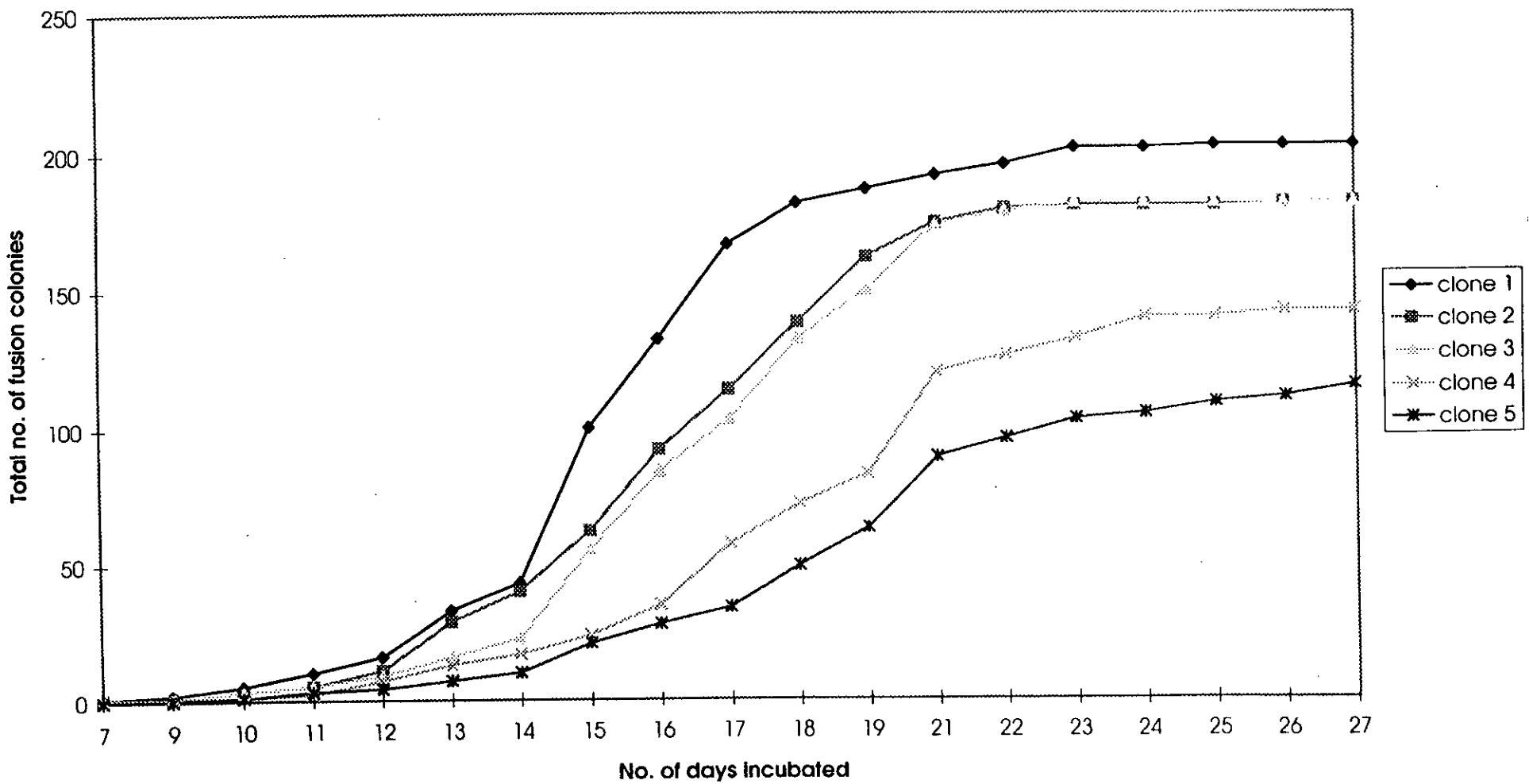


Figure 35. Ara-Lac fusion colony formation in DL943 sbcCD

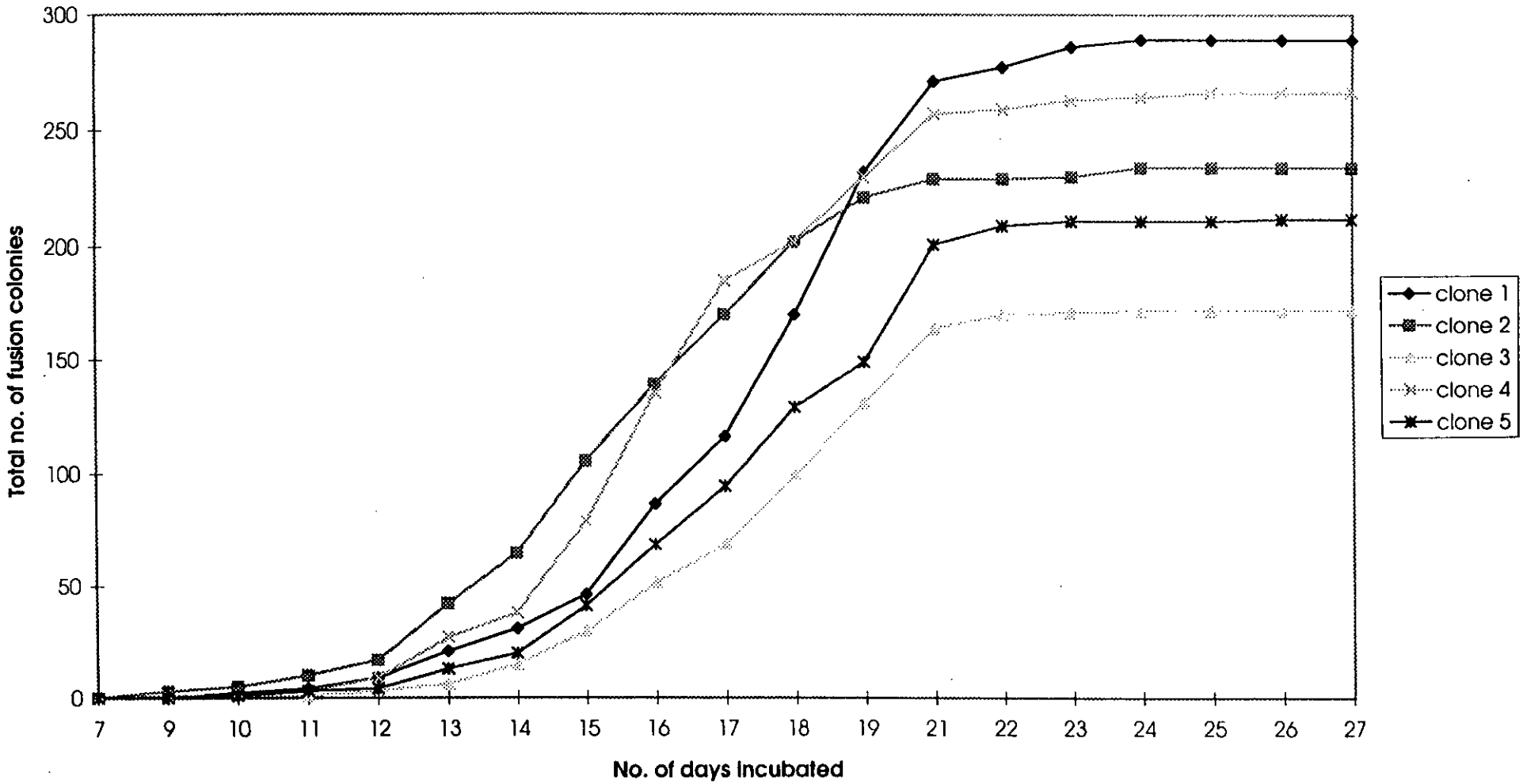


Figure 36. Ara-Lac fusion colony formation in DL944 mutS

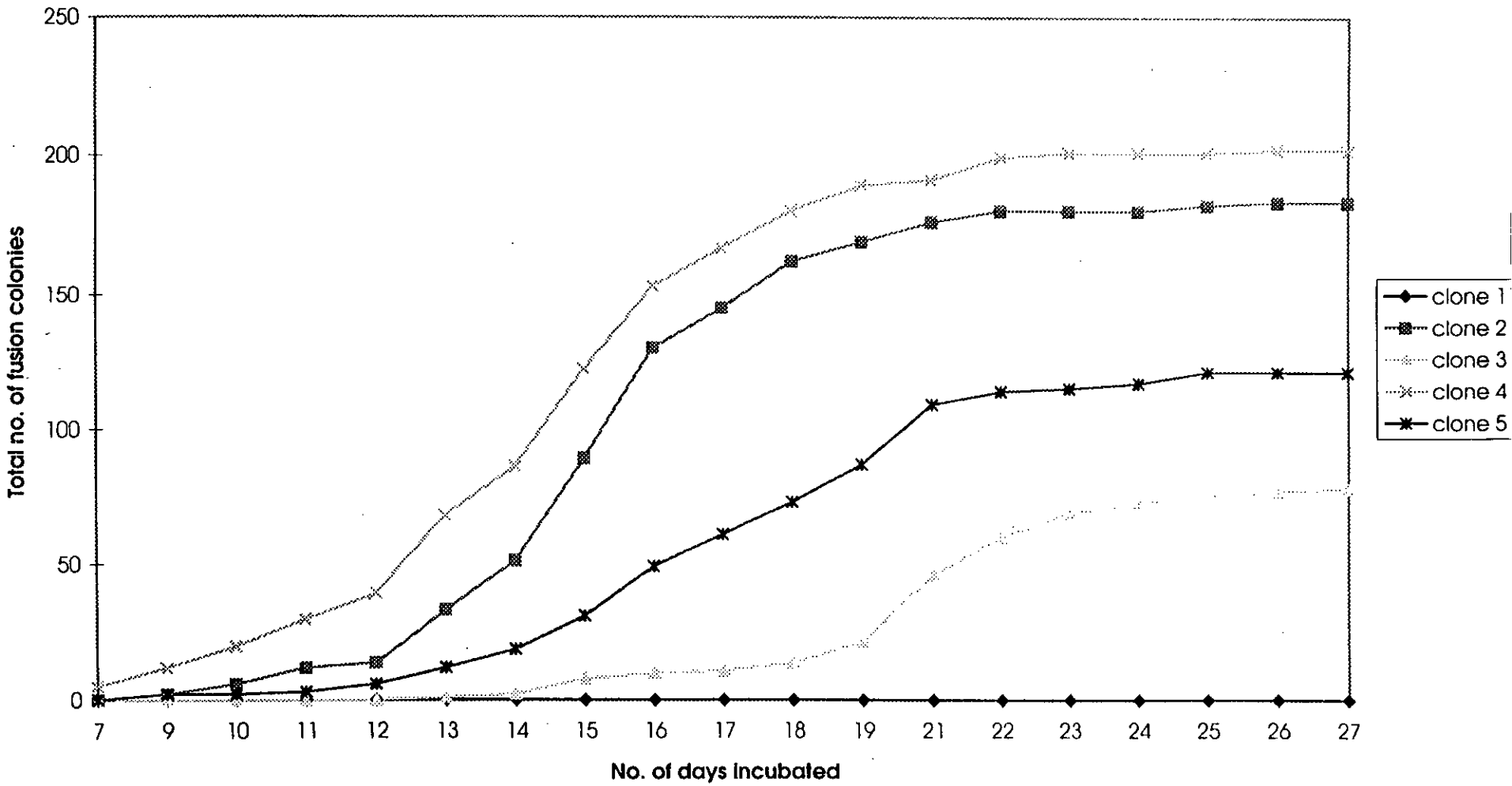
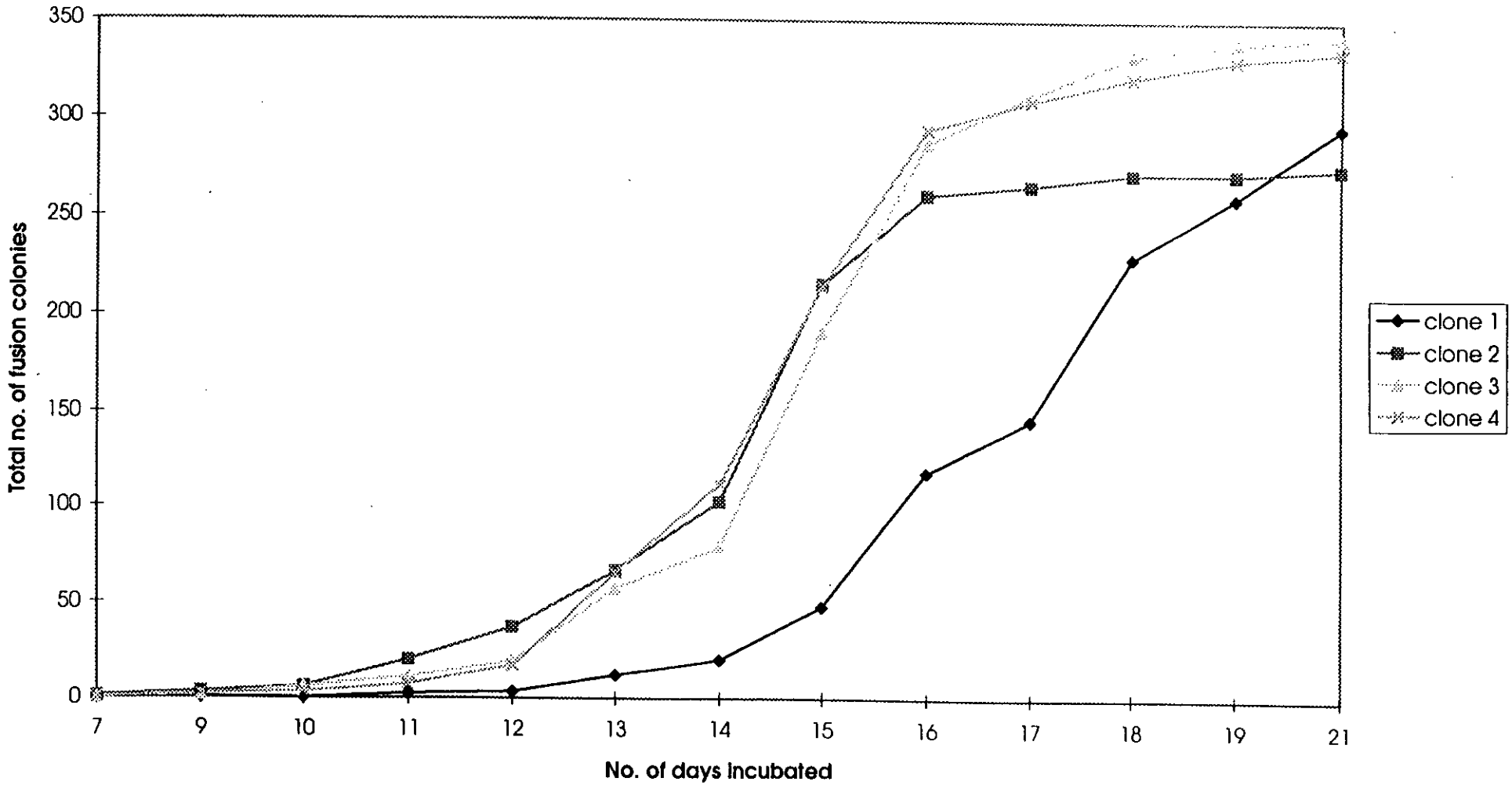


Figure 37. Ara-Lac fusion colony formation in DL945 sbcCD mutS

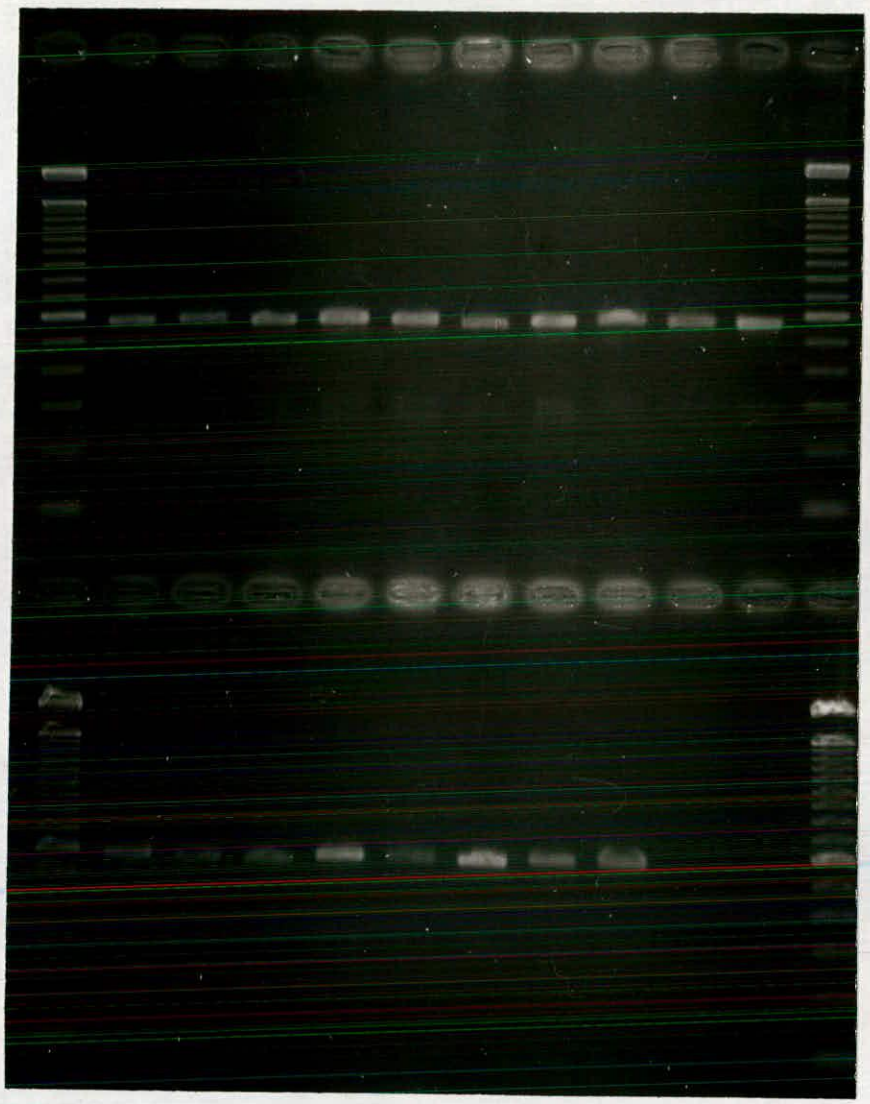


It has previously been reported that the only requirement for fusion formation are starvation conditions. Experiments were performed by Mittler and Lenski (1990) and Maenhault-Michel and Shapiro (1994) in which fusion strains were grown in aerated and unaerated glucose starved cultures. Plating out of the cultures over time resulted in the appearance of fusion colonies on medium containing arabinose and lactose within two days after plating. Colonies which appeared on the plates within two days were assumed to have formed within the initial starved culture. The starvation experiment was repeated using the strains constructed in this study however no fusion colonies were detected by this method.

Fusion colonies were purified from plates for the wild-type (DL942), *sbcCD* (DL943), *mutS* (DL944) and *sbcCD mutS* (DL945) and the region around the fusion junction amplified by PCR using primers previously described (Maenhault-Michel and Shapiro 1994). The DNA fragments amplified were electrophoresed on 2% NuSieve agarose gels (FMC Bioproducts) and the sizes determined to be between 0.5 and 0.65 kb. No differences were seen in the sizes of PCR fragments derived from fusion colonies produced by strains bearing chromosomal mutations and the wild-type strain (figures 38, 39 and 40). There was also little variation in size of PCR products produced from colonies which form on selective medium as previously reported (Maenhault-Michel and Shapiro, 1994).

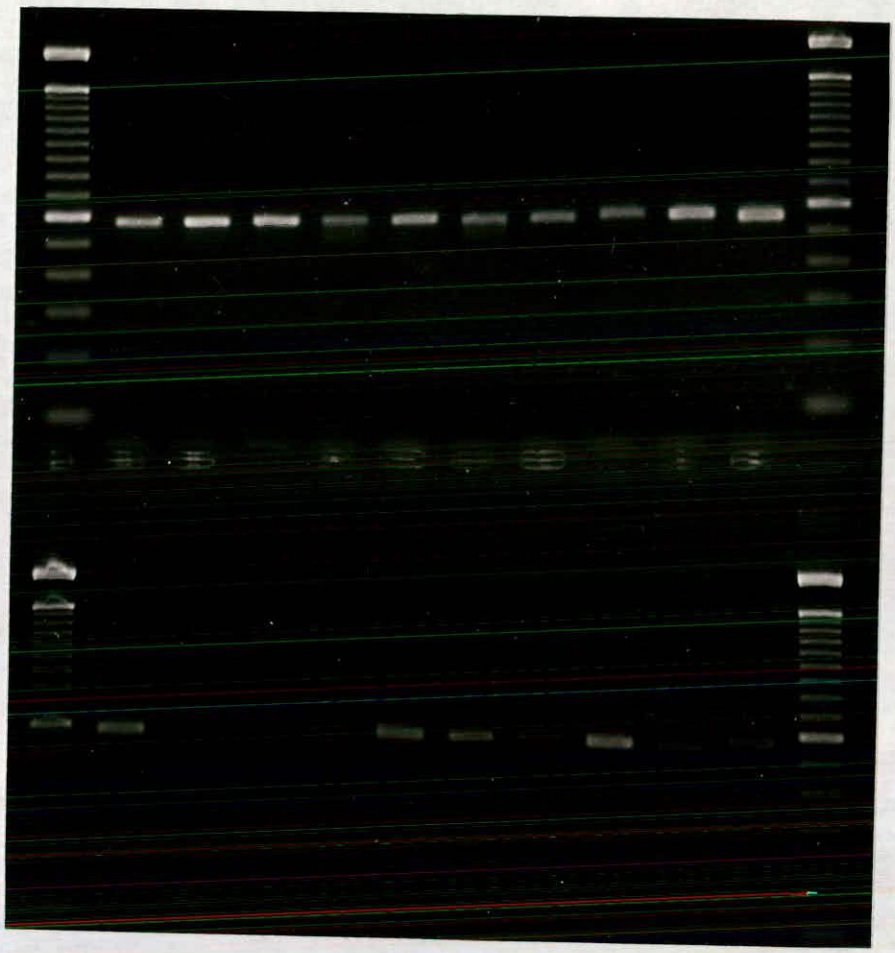
A total of 31 PCR products derived from fusion colonies from the different backgrounds used in this study were sequenced (figure41). Although no effect of the chromosomal mutations on the kinetics of fusion formation was detected, the fusion junctions were sequenced to see if the mutations affected the fusion process in other more subtle ways. All of the fusions had a portion of *Mu* R adjacent to *lacZ* . The presence of a *MuR-lacZ* junction indicated a role for *Mu* transposition functions (Shapiro and Leach, 1990). Of the 31 PCR products sequenced, four different target sites for *Mu* were identified in *lacZ* (figure 42). Of these 4 target sites, 1 accounts for 21 of the 31 fusions isolated (junction 2 ~68%) and matched the NYG/CRN consensus of the *Mu* target site determined *in vivo* (Mizuuchi and Mizuuchi, 1993). This target site is the same as that determined for the previously published fusion sequence (Maenhault-Michel and Shapiro, 1994). The 3 remaining junctions accounted for 10 out of 31 of the PCR products (~32%). Two different *araB-Mu* junctions were identified (figure 42) and one of these junctions (junction A) was found in 30 out of the 31 fusions sequenced (~97%).

Figure 38. PCR analysis of *araB-lacZ* fusion junctions in a wild-type host.



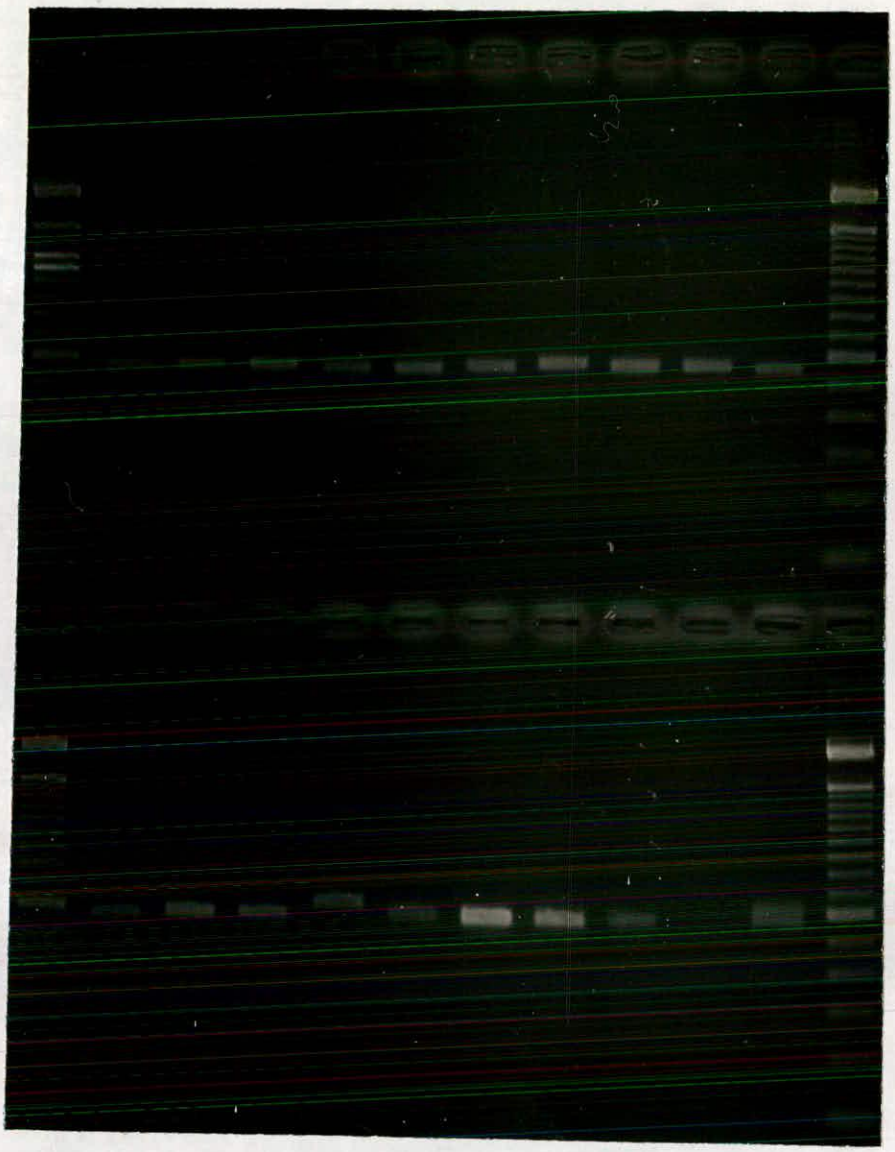
Amplification carried out using primers previously described (Maenhault-Michel and Shapiro, 1994). PCR products electrophoresed on 2% NuSieve agarose gels (FMC Bioproducts). 100 bp size marker ladder present in outside wells (Gibco BRL).

Figure 39. PCR analysis of *araB-lacZ* fusion junctions in an *sbcCD* host.



Amplification carried out using primers previously described (Maenhault-Michel and Shapiro, 1994). PCR products electrophoresed on 2% NuSieve agarose gels (FMC Bioproducts). 100 bp size marker ladder present in outside wells (Gibco BRL).

Figure 40. PCR analysis of *araB-lacZ* junctions in a *mutS* and *sbcCD mutS* hosts.



Amplification carried out using primers previously described (Maenhault-Michel and Shapiro, 1994). PCR products electrophoresed on 2% NuSieve agarose gels (FMC Bioproducts). 100 bp size marker ladder present in outside wells (Gibco BRL).

Figure 41.

Sequences of *araB-lacZ* fusions. Obtained by direct sequencing of PCR products using Sequenase PCR product sequencing kit (USB). Ambiguities present within some of the sequences may be due to PCR artefacts or the presence of more than one PCR product within the sequencing reaction.

Figure 41. Sequence of *araB-lacZ* fusion junctions.

Wild-type fusions	No. sequenced
GGTTT↓TTCGTGCGCCGCTTCA↑CTGGCGTTACCCAACTTA	3
GGTTT↓TTTCGTGCGCCGCTTCA↑CCCAACTTAATCGCCTT	2
GGTTT↓TTTCGTGCGCCGCTTCA↑GCCTTGCAGCACATC	1
GGTTTAACGGTCGT↓CGCTTCA↑CCCAACTTA	1
***SbcCD* fusions**	
GGTTT↓TTCGTGCGCCGCTTCA↑CTGGCGTTACCCAACTTA	3
GGTTT↓TTCGTGCGCCGCTTCA↑ATCGCCTTGCAGTCACAT	1
***MutS* fusions**	
GGTTT↓TTCGTGCGCCGCTTCA↑CTGGCGTTACCCAACTTA	1
GGTTT↓TTCGTGCGGCGTTCA↑CTGGCGTTACCCTAACTTA	10
GGTTT↓TTCGTGCGCCGCTTCA↑CCCAACTTAATCTGA	2
GGTTT↓TTCGTGCGCCGCTTCA↑ATCGCCTTGCAGTCACA	2
***SbcCD MutS* fusions**	
GGTTT↓TTCGTGCGCCGCTTCA↑CTGGCGTTACCCTAACTT	4
GGTTT↓TCGTGCGCCGCTTCA↑ATCGCCTTGCAGTCACAT	1

Sequences derived from direct sequencing of PCR products.

T represents ambiguities in sequence (either insertions or substitutions).

↓ *araB-Mu*R fusion junction.

↑*lacZ-Mu*R fusion junction.

Figure 42. Sequences of *araB-Mu* and *Mu-lacZ* junctions.

A. *AraB- Mu*R Junctions

TGCTCGACTG**GTTT**TTCGTGCGCCGCTTCA JUNCTION A

TGCTCGACTGGTTTAACGGT**CGT**CGCTTCA JUNCTION C

B. *LacZ-Mu*R junctions

CTTCA**CTGGCGTT** JUNCTION 2

CTTCA**CCCAACTT** JUNCTION 3

CTTCA**ATCGCCTT** JUNCTION 5

CTTCA**CCTTGCAG** JUNCTION 6

40 a. Junctions of *araB-Mu* fusions. Arrows show orientation of short repeated sequence.

40 b. Sequence shown in bold is that corresponding to *lacZ*.

Discussion

Shapiro and Leach (1990) proposed a strand transfer model to account for the involvement of *Mu* transposition in fusion formation. The first step in the fusion process involves the formation of a plectonemic complex which brings together the two ends of the *Mu* genome and the target sequence present in the 5' end of the *lacZ* gene. A strand transfer reaction takes place which results in the ligation of the 3' hydroxyl groups from the *Mu* ends and the 5' phosphate groups from the *lacZ* target sequence. The ligation reaction leaves two exposed 3' hydroxyl groups in the *lacZ* target sequence which could act as primers for chain elongation during leading strand replication. Replication could proceed into the end of the *Mu* molecule. Strand switching of the nascent strand or a cutting and patching reaction at the end of the *Mu* followed by continued replication into *araB* would result in a molecule composed of *araB-Mu*R-*lacZ*. This model accounts for all of the sequences determined in this study.

As discussed in Maenhault-Michel *et al* (1996) the crossover points for fusion formation occur within short directly repeated sequences present within *araB* and *Mu* (figure 43). Formation of the strand transfer complex described by Shapiro and Leach (1990) followed by isomerisation to a four way junction places the first nucleotide of both of the short direct repeats at the branch point of the molecule (figure 43). This branched molecule could be subject to one of three fates. Firstly,

cleavage across the junction followed by ligation of the free ends, similar to the Holliday junction resolution activity of RuvABC (West, 1994). Secondly, enzymatic degradation of two arms of the four way junction from single strand regions present at mismatches followed by ligation of the free DNA ends. Thirdly, replication through *lacZ* into *Mu*R followed by strand slippage to a short homologous sequence in *araB*. A second round of replication would effectively fix the mutation allowing transcription of *lacZ* and removal of the strand containing the branched structure. Although there was no apparent effect of an *sbcCD* mutation on either the kinetics of fusion formation or the sequences of the fusions formed, SbcCD might have a role in the degradation or processing of the intermediate which is not detected in this assay.

Figure 43.

Phage *Mu* inserts into the chosen cistron. Recombination then occurs the inserted *Mu* and a terminal fragment of *Mu* located upstream of a decapitated *lacZ* gene carried on a λ*plac* bacteriophage. The *lacZ* gene is now positioned downstream of *Mu*. There is no promoter for *lacZ* transcription and the presence of an ochre triplet at codon 18 blocks all transcription and translation of the *lacZ* gene. Excision of the *Mu* sequence can lead to the in-frame fusion of *araB* and *lacZ* and subsequent transcription and translation.

Figure 43. Construction of fusion strain MCS2a2 (after Casadaban, 1976)

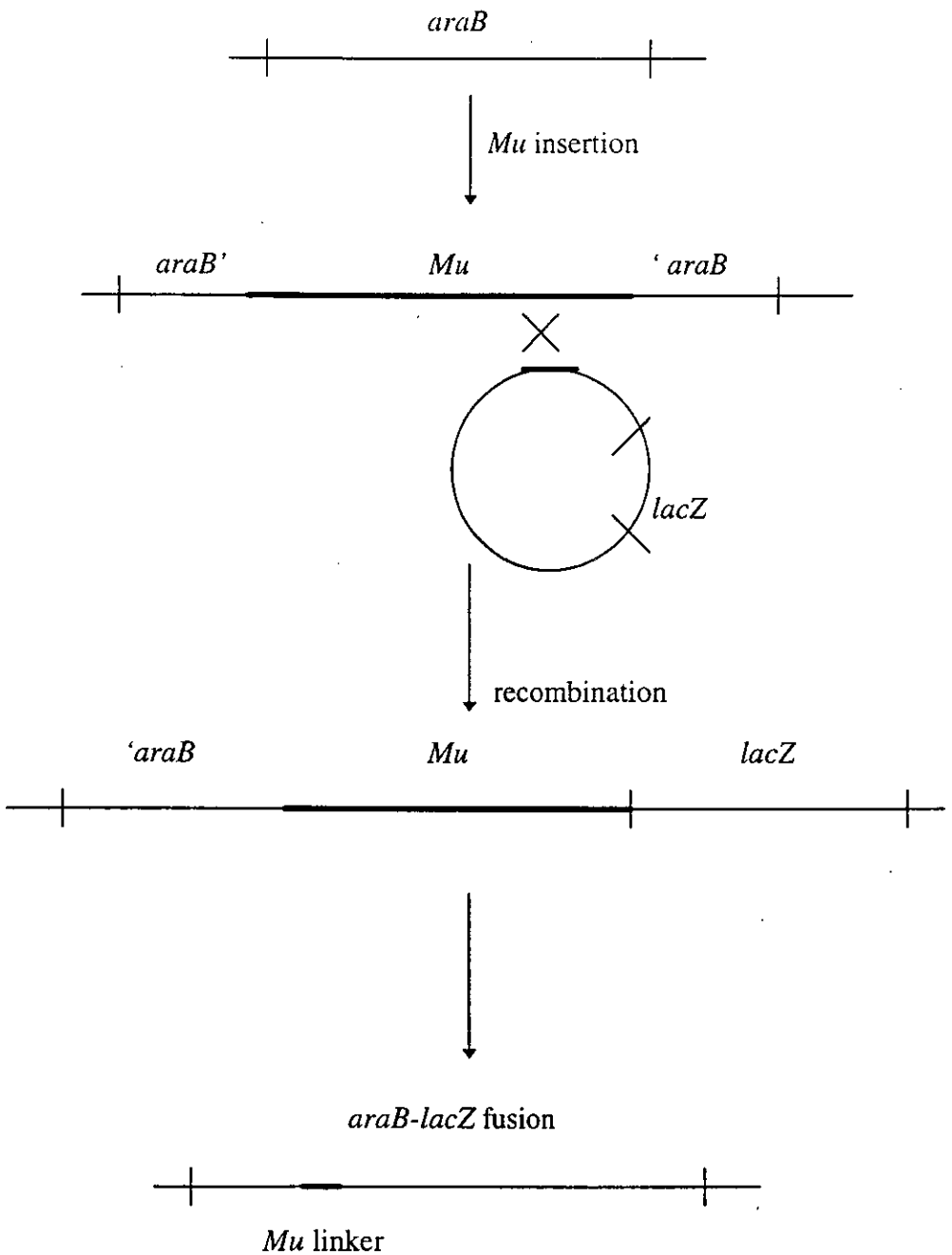
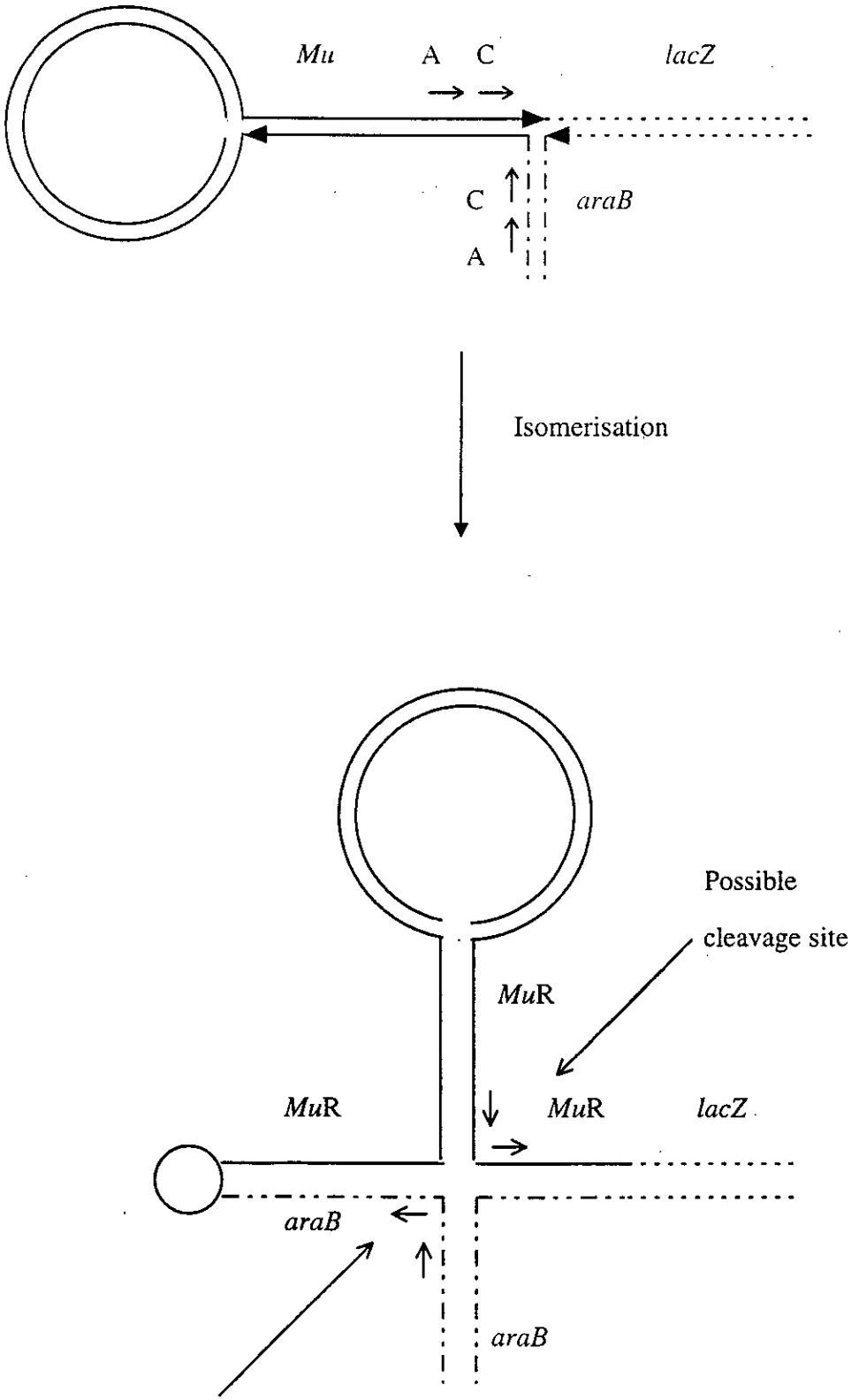


Figure 44.

Strand transfer complex showing orientation of short direct repeats present within *araB* and *Mu*R. Isomerization of the strand transfer complex gives rise to a four way junction with the short direct repeats placed at the branch points of the structure. The four way junction could be subject to cleavage across the junction to give rise to the fusions detected by the assay.

⟶ Short repeats present within *araB* and *Mu*R

Figure 44. Strand transfer complex showing position of short direct repeats.



CHAPTER 6
DISCUSSION

Discussion

In this work the deletion of long DNA palindromes from high copy number plasmids was examined. It was demonstrated that a 571 bp λ-derived palindrome apparently deleted from pAC2 (pUC18 +571 bp palindrome) with a frequency of ~50%. Plating of cells bearing the plasmid on selective medium containing a chromogenic substrate gave rise to blue/white sectored colonies. Subsequent work by Pinder *et al.* (1996, submitted for publication) revealed that the apparent sectoring phenotype of cells containing plasmid pAC2 was due to its structure as a heterodimer made up of one copy of the palindrome and a second copy a short repeated sequence of ~40 bp of the plasmid multiple cloning site. It was the presence or instability of this short repeated sequence which gave rise to the sectoring colonies, although deletion of the remaining copy of the palindrome may also have contributed to the sectoring.

Plasmid derivatives of pAC2 (pMS5 and pMS7) which appeared to allow the stable maintenance of long palindromes (Shaw and Leach, unpublished) were investigated to determine the nature of the stability. The palindromes deleted from the plasmid with a frequency of ~2% for pMS5 and <0.0025% for pMS7. Deletion was mediated by 3 bp direct repeats. A number of molecular techniques were used to investigate the source of the increased stability including restriction analysis, DNA sequencing and estimation of supercoiling level. No mutation could be found which

could lead to the increased palindrome stability, although it appeared that pMS7 was less negatively supercoiled than pMS5.

E. coli strains which carry mutations in the *sbcC* and/or *sbcD* gene are used for the propagation of long palindromes in λ cloning vectors since they overcome the inviablity associated with this type of sequence (Leach *et al.* 1987). In this work a plasmid bearing a long DNA palindrome could also be propagated in a plasmid when when grown in an *sbcCD* strain. When plasmid DNA bearing the 571 bp palindrome was prepared from strains which were deleted for the *sbcC* and *sbcD* genes no plasmid monomers were detected for pAC2 and pMS7 and only ~4% of the total DNA content was present as monomers for pMS5. The palindrome-free plasmid pUC18 could be isolated as monomers which accounted for approximately 30% of the total DNA content. It appears that the presence of the palindrome in the plasmid leads to increased mulitmerisation of the plasmid DNA. The presence of long palindromes may promote RecA-dependent interplasmidic recombination and so give rise to large amounts of multimers. When artificially created plasmid monomers of pAC2, pMS7 and pUC18 were produced and propagated in a *recA sbcCD* background only one colony was isolated for pMS7. This single colony contained pMS7 DNA in a monomeric form and from the size of this plasmid it was assumed that a full length palindrome was still present. However, since only one colony could be isolated it is possible that the plasmid may have been a variant of the original pMS7, perhaps bearing a mutation to allow propagation in a monomeric form. The

colonies isolated for pAC2 contained monomeric DNA which did not contain a palindrome but apppeared to be larger than monomers of pUC18 indicating the presence of the additional sequence described by Pinder *et al.* (submitted for publication). From this work it is apparent that plasmids bearing long palindromes are most easily propagated in a multimeric form.

A short 109 bp palindrome was cloned into the related plasmid pUC118 and pUC119 to give rise to the plasmid pDLJ1 and pDLJ2 (D. Leach, unpublished). The palindrome deleted from the plasmids by an illegitmate recombination event involving 7 bp direct repeated sequences. The position of the repeats with respect to the direction of replication and the frequency of deletion in each plasmid lead to the conclusion that deletion occurred preferentially on the lagging strand of the replication fork. This observation is in agreement with earlier work performed by Trinh and Sinden (1991) who used a different experimental system. However, Weston-Hafer and Berg (1990) could find no difference in the frequency of deletion of short palindromic sequences from either the lagging or leading strand of the replication fork. It should be noted that in the work presented here and in that performed by Trinh and Sinden (1991), the region of DNA which contains the palindrome and direct repeats is inverted to switch the preferred position of the repeats from the leading to the lagging strand. In this way the local sequence context of the palindromes and the repeats remains unchanged. In the work by Weston-Hafer and Berg (1990) large regions of the plasmid DNA containing the plasmid origin of

replication were inverted, possibly leading to changes in the tertiary structure of the plasmid. Inversion of the replication of origin also changed the direction of replication with respect to the direction of transcription through the *tet* gene containing the palindromic sequences. Given the influence of local and global sequence context on deletion formation (see Chapter 1) it is possible that any influence of replication direction could have been masked by rearrangement of the plasmid structure.

Plasmid pDLJ2 (bearing the 109 bp palindrome) also forms multimers in an *sbcCD* background although some monomers are present. Unlike the 571 bp palindrome which could not be propagated in a wild-type background (because of the associated inviability) pDLJ2 can be grown in a wild-type host. When pDLJ2 is propagated in a wild-type background, only plasmid monomers are present which is surprising since it might be expected that the presence of RecA could promote interplasmidic recombination and so give rise to plasmid multimers (which is observed for the palindrome free plasmid pUC18).

The presence of a functional SbcCD protein in the wild-type strain may lead to breakage of a replication fork because of action of the SbcCD protein on a hairpin or cruciform structure formed by the palindrome. Breakage of the replication fork gives rise to a σ structure leading to rolling circle replication. The double strand tail of the rolling cirlce molecule would be open to attack by RecBCD nuclease leading

to decreased yields of plasmid DNA in a wild-type host, which is observed in this work. Recombination between monomeric units of the plasmid produced by rolling circle replication would lead to the production of purely monomeric DNA. In an *sbcCD* background any secondary structures formed by the palindromes are not subject to processing by SbcCD so the is little or no breakage of the replication forks. As a result plasmid multimers can accumulate which increase the stability of palindromes as seen with the 571 bp palindrome.

Warren and Green (1985) investigated plasmid pBR322 bearing palindromes which were 147 and 146 bp in length. It appeared that the presence of the palindromes inhibited multimer formation in a *recA*$^{+}$*sbcB*$^{-}$ background, despite earlier evidence that this genetic background enhanced multimer formation. A shorter 114 bp palindrome did not exhibit this inhibition. The results of their study are shown in the following table:-

Plasmid	**Monomers %**	**Dimers %**	**Multimers %**
pBR322	~84.0	14.8	<1.2
114 bp pal	79.0	19.4	1.6
146 bp pal	~93.9	4.9	<1.2
147 bp pal	~98.3	0.9	<0.8

The *sbcC* mutation was originally isolated as a cosuppressor of *recBC* in *sbcB* backgrounds (Lloyd and Buckman, 1985). It seems unlikely that there is an additional *sbcC* mutation present in the *E. coli* K-12 strain used in the study by Warren and Green (1985) since cloning of the longer palindromes resulted in a three-fold reduction in plasmid copy number whereas no effect was seen for the 114 bp palindrome. This may be indicative of palindrome inviability in this background. The work by Warren and Green (1985) supports the idea that there is an inhibition of multimerisation by palindrome bearing plasmids in an *sbcCD*$^+$ background.

From the work performed here and that of Pinder *et al.* (submitted for publication) it is apparent that plasmid pAC2 exists as a heterodimer, with one half of the dimer bearing a mutation in the form of a repeated sequence of ~40 bp length. It seems likely that pMS7 is in fact the original form of the plasmid and that pAC2 is a derivative of the pMS7 arising as a result of aberrant replication of one copy of the palindrome. Experiments performed by other workers have looked at the influence of plasmid multimerisation and heterodimer formation on the fixing of mutations. Studies carried out by Dianov *et al* (1991) and Mazin *et al.* (1991) looked at the deletion of long (165 bp and 401 bp) and short (13, 21 and 42 bp) direct repeats respectively. The repeats were cloned into the *tet* gene of plasmid pBR322 and deletion of the repeats should restore tetracycline resistance. Deletion of the 401 bp repeat was 2.5-fold higher than the 165 bp repeat in an *E. coli* AB1157 wild-type background. In a *recA* background the deletion frequency was decreased 8-fold for

the 401 bp repeats and 10-fold for the 165 bp repeats. Deletion was also reduced in a *recF, recJ* or *recO* background. When the DNA from colonies which had reverted to tetracycline resistance was examined it was found that the DNA was only present as dimers and monomers were never isolated. DNA in the tetracycline sensitive colonies which had not undergone deletion was present only as monomers.

When the structure of the tetracycline genes was investigated two structurally different types of plasmid dimer were identified. Each dimer had two copies of the *tet* gene, one of which was a wild-type copy , the other copy containing a repeated sequence. The repeated sequence was either made up of the original 165 or 401 bp repeats (this was produced by a RecA-independent mechanism involving short direct repeats) or three copies of the repeats (produced by interplasmidic recombination involving unequal crossing-over in a RecA-dependent mechanism). In a wild-type background dimers bearing three repeats dominated amongst the plasmid population (~85-90%). Deletion of the short direct repeats used in the study by Mazin*et al.* (1991) was again associated with plasmid dimerization. All the *tet* ^{R} colonies examined contained dimeric DNA which was predominantly of the structure which contained one wild-type copy of the *tet* gene and one copy containing the original number of repeats. Based on these studies Mazin *et al.* (1996) proposed that under conditions of selection, dimers provide a mechanism for the rapid accumulation of advantageous mutations since dimers accelerate the segregation of revertant and parental plasmids.

The formation of dimers decreases the copy number of the plasmid since copy number control systems count the total number of plasmid origins within the cell and not the number of plasmid present; this leads to an increase in the number of plasmid free cells (Summers and Sherratt, 1884). The formation of plasmid free cells would account for the poor colony morphology seen with pDLJ2 in an *sbcCD* background. These plasmid free cells would not survive on the selective media used. The introduction of a *recA* mutation into the *sbcCD* strain alleviated the poor colony morphology and lead to the formation of plasmid monomers indicating an effect of plasmid multimerisation on the morphology of the cell. Another consequence of plasmid dimerization is the high physiological load place on the cell which results in slower growth than monomer containing cells (Summers *et al.*, 1993). However, this is counteracted by an increase in the replication rate since each plasmid dimer contains two replication origins. In this way dimers would be favoured over monomers because they are capable of faster replication. The increased replication may increase the chances of a palindromic sequence being replicated successfully and so there would be a selective advantage in being in a dimeric state.

The work presented in this thesis has direct relevance to the choice of host when cloning long palindromic sequences which are present within eukaryotic genomes. A 571 bp palindrome could be successfully cloned into plasmid pUC18 to give the plasmid pMS7 (Shaw and Leach, unpublished). The plasmid could propagated in an *sbcCD* background with low levels of instability. This stability

seems to be dependent on the plasmid DNA being in a multimeric form since it was difficult to produce monomers of this plasmid and propagate them in a *recA sbcCD* strain. A shorter 109 bp palindrome could be propagated in all strains examined however, in a wild-type background DNA yields were low compared with an *sbcCD* strain and the *sbcCD* strain was subject to poor colony morphology possibly because of the production of plasmid free cells as a result of plasmid multimer production. The use of a *recA sbcCD* strain resulted in the formation of plasmid monomers and good colony morphology. From this work it is suggested that an *sbcCD* strain be used for the cloning of palindromic sequences in plasmids, although the effect on plasmid multimerisation of the host strain should be taken into account. In some cases it may be preferable to use a *recA sbcCD* strain.

Using flow cytometry an asynchrony phenotype is seen with a *recA* host in which partial chromosomes are present within the cell (Skarstad and Boye, 1993). It was proposed that the asynchrony was abolished in a *recD* background because the RecBCD nuclease was responsible for the degradation of collapsed replication forks. The effect of an *sbcCD* mutation on the asynchrony phenotype was examined to see if there was a role for the SbcCD protein in the degradation of broken replication forks. No effect was seen perhaps indicating that other more powerful nucleases are at work within the cell or that there may be alternative ways of breaking a replication fork other than by the effect of SbcCD on an unusual secondary structure.

Finally the effect of SbcCD on the formation of *araB-lacZ* mutations was investigated. Using techniques previously described (Maenhault-Michel and Shapiro 1994) it was shown that an *sbcCD* mutation did not influence the kinetics of fusion formation or the sequences of the fusions points investigated. However, given the complex structure of the intermediate involved in fusion formation the SbcCD protein may have a role in the processing of the structure which does not effect the sequence of the fusions detected.

CHAPTER 7
REFERENCES

Albertini, A.M., Hofer, M., Calos, M.P. & Miller, J.H. On the formation of spontaneous deletions: The importance of short sequence homologies in the generation of large deletions.

Cell **29**: 319-328 (1982)

Allgood, N.D. & Silhavy, T.J. Illegitimate recombination in bacteria.

In: Genetic Recombination, pp. 309-330 (1988) (American Society for Microbiology, Washington, DC).

Eds. Kucherlapati, R & Smith, G.R.

Allgood, N.D. & Silhavy, T.J. *Escherichia coli xonA* (*sbcB*) muatnts enhance illegitimate recombination.

Genetics **127**: 671-680 (1991)

Anderson, C.W. Sponteous deltion formation in several classes of *Escherichia coli* mutants deficient in recombination ability.

Mutation Res. **9**: 155-165 (1970)

Anderson, R.P., & Roth, J. R. Tandem genetic duplications in phage and bacteria.

Annu. Rev. Microbiol. **31**: 473-505 (1977)

Anderson, P. & Roth, J. Spontaneous tandem genetic duplications in *Salmonella typhimurium* arise by unequal recombination between rRNA (rrn) cistrons.

Proc. Natl. Acad. Sci. USA **78**: 3113-3117 (1981)

Arber, W. Natural mechanisms of microbial evolution

In: Genetic Manipulation: Impact on Man and Society, pp. 1-14 (1984) (Cambridge University Press, Cambridge)

Ed. W.Arber

Bachellier, S., Saurin, W., Perrin, D., Hofnung, M. & Gilson, E. Structural and functional diversity among bacterial interspersed mosaic elements (BIMEs).

Mol. Microbiol. **12**: 61-70 (1994)

Bachman, B.J. Derivations and genotypes of some mutant derivatives of *Esherichia coli* K-12

In: *Escherichia coli* and *Salmonella typhimurium* Vol. 2, pp. 1190-1219 (1987) (American Society for Microbiology, Washington, D.C.)

Eds. Neidhardt, F.C, Ingraham, L.J., Low, K.B., Magasnik, B. & Schaechter, M.

Behnke, D., Malke, H., Hartman, M. & Walter, F. Post-transformational rearrangement of an *in vitro* reconstructed group-A streptococcal erythromycin resistance plasmid.

Plasmid 2: 605-616 (1979)

Benson, S.A. & Bremer, E. In vivo selection and characterization of internal deletions in the *lamB::lacZ* gene fusion.

Gene **52**: 165-173 (1987)

Berg, D.E., Egner, C., Hirschel, B.J., Johnsrud, L., Jorgensen, R.A. & Tlsty, T.D. Insertion, excision, and inversion of Tn5.

Cold Spring Harbor Symp. Quant. Biol. **45**: 115-123 (1980)

Berg, D.E., Egner, C. & Lowe, J.B. Mechanism of F factor-enhanced excision of tranposon Tn5

Gene **22**: 1-7

Berg, D.E. Transposon Tn5

In: Mobile DNA, pp. 185-210 (1989) (American Society for Microbiology, Washington, D.C.)

Eds. Berg, D.E. & Howe, M.M.

Bertrand, K., Korn, L., Lee, F., Platt, T. Squires, C.L. Squires, C. & Yanofsky, C. New features of the regulation of the tryptophan operon.

Science **189**: 22-26 (1975)

Birnboim, H.C. & Doly, J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA.

Nucleic Acids Res. **7**: 1513-1523 (1979)

Brunier, D., Peeters, B.P.H., Bron, S. & Ehrlich, S.D. Breakage-reunion and copy choice mechanisms of recombination between short homologous sequences.

EMBO J. **8**: 3127-3133 (1989)

Burland, V., Plunkett, G., Daniels, D.L. & Blattner, F.R. DNA sequence and analysis of 136 kilobases of the *Escherichia coli* genome: Organizational symmetry around the origin of replication

Genomics **16**: 551-561

Campbell, A. Genetic recombination between λ prophage and irradiated λ*dg* phage.

Virology **23**: 234 (1964)

Casadaban, M.J. & Cohen, S.N. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli.*

J. Mol. Biol. **138**: 179-207 (1980)

Chalker, A.F. *SbcC* and palindrome-mediated inviability in *Escherichia coli.*

Ph.D. Thesis, 174 pp. (1990) (University of Edinburgh)

Chalker, A.F., Leach, D.R.F. & Lloyd, R.G. *Escherichia coli sbcC* mutants permit stable propagation of DNA replicons containing a long palindrome.

Gene (Amst.) **71**: 201-205 (1988)

Champoux, J.J. & Dulbecco, R. An activity from mammalian cells that untwists superheical DNA - a possible swivel for DNA replication.

Proc. Natl. Acad. Sci. USA **69**: 143-146 (1972)

Chedin, F., Dervyn, E., Ehrlich, S.D. & Noirot, P. Frequency of deletion formation decreases exponentially with the distance between short direct repeats.

Molecular Microbiology **12**: 561-569 (1994)

Collins, J. Instability of palindromic DNA in *Escherichia coli.*

Cold Spring Harbor Symp. Quant. Biol. **45**: 409-416 (1981)

Collins, F. S. & Weissman, S.M The molecular genetics of human haemoglobin.
Nucleic Acids Res. **31**: 315-436 (1984)

Connelly, J. & Leach, D.R.F The SbcCD nuclease of *Escherichia coli.*
J. Cell. Biochem. **19**: 102 (1995)

Connolly, B., Parsons, C.A. & Benson, F.E. Resolution of Holliday junctions *in vitro* requires the *Escherichia coli rUVC* gene product.
Proc. Natl. Acad. Sci. USA **88**: 6063-6067

Coukell, M.B. & Yanofsky, C. Increased frequency of deletions in DNA polymerase mutants of *Escherichia coli.*
Nature (Lond.) **228**: 633-635 (1970)

Croce, C.M. Role of chromosome translocations in human neoplasia.
Cell **49**: 155-156 (1987)

d'Alencon, E., Petranovic, M., Michel, B., Noirot, P. Aucouturier, A., Uzest, M. & Ehrlich, S.D. Copy-choice illegitimate DNA recombination revisited.
EMBO J. **13**: 2725-2734 (1994)

Dianov, G.L., Kuzminov, A.V., Mazin, A.V. & Salganik, R.I. Molecular mechanisms of deletion formation in *Escherichia coli* plasmids. 1. Deletion formation mediated by long direct repeats.
Mol. Gen. Genet **228**: 153-159 (1991)

DasGupta, U., Weston-Hafer, K. & Berg, D.E. Local DNA sequence control of deletion formation in *Escherichia coli* plasmid pBR322.
Genetics **115**: 441-49 (1987)

DiGate, R.J. & Mariens, J. Molecular cloning and DNA sequence analysis of *Escherichia coli topB*, the gene encoding topoisomerase III.
J. Biol. Chem. **264**: 17924-17930 (1989)

Donachie, W. What is the minimum number of dedicated functions required for a basic cell cycle?

Curr. Opin. Genet. and Dev. **2**: 792-798 (1992)

Drlica, K. Biology of bacterial deoxyribonucleic acid topoisomerases

Microbiol. Rev. **48**: 273-289 (1984)

Edlund, T. & Normark, S. Recombination between short DNA homologies causes tandem duplication.

Nature **292**: 269-271 (1981)

Egner, C. & Berg, D.E. Excision of transposon Tn5 is depenedent on the inverted repeats but not on the transposase function of Tn5.

Proc. Natl. Acad. Sci. USA **78**: 459-463 (1981)

Ehrlich, S.D. Illegitimate recombination in bacteria.

In: Mobile DNA, pp. 799-832 (1989) (American Society for Microbiology, Washington, D.C.)

Eds. Berg, D.E. & Howe, M.M.

Ehrlich, S.D., Bierne, H. & d'Alencon, E. Mechanisms of illegitimate recombination.

Gene **135**: 161-166 (1993)

Fairfield, F.R., Bauer, W.R. & Simpson, M.V. Mitochondria contain a distinct DNA topoisomerase.

J. Biol. Chem. **254**: 9352-9354 (1979)

Farabaugh, P.J., Schmeissner, U., Hofer, M. & Miller, J.H. Genetic studies of the *lac* repressor VII. On the molecular nature of spontaneous hotspots in the *lacI* gene of *Escherichia coli.*

J. Mol. Biol. **126**: 847-863 (1978)

Forrest, S.M., Cross, G.S., Speer, A., Gardner-Medwin, D., Burn, J. & Davies, K.E. Preferential deletion of exons in Duchenne and Becker muscular dystrophies.

Nature **329**: 638-640 (1987)

Foster, P.L. Directed mutation: between unicorns and goats.

J. Bacteriol. **174**: 1711-1716 (1992)

Foster, P.L. Adaptive mutation: the uses of adversity.

Ann. Rev. Microbiol. **47**: 467-504 (1993)

Foster, T.J., Lunblad, V., Hanley-Way, S. Halling, S.M. & Kleckner,N. Three Tn10-associated excision events: relationship to transposition and role of direct and inverted repeats.

Cell **23**: 215-227 (1981)

Franklin, N.C. Extraordinary recombinational events in *Escherichia coli.* Their independence of the *rec*$^+$ function.

Genetics **55**: 699-707 (1967)

Franklin, N.C. Illegitimate recombination.

In.: The Bacteriophage Lambda, pp. 175-194 (1971) Cold Spring Harbor Monograph Series (Cold Spring Harbor Laaboratory Press, Cold Spring Harbor, New York).

Ed. Hershey, A.D.

Friedberg, E.C. DNA Repair.

W.H. Freeman and Co. San Franscico (1992)

Friend, S.H., Bernards, R. Rogelj, S., Weinberg, R.A, Rapaport, J.M., Albert, D.M. & Dryja, T.P. A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma.

Nature **323**: 643-646 (1986)

Gellert, N. DNA topoisomerases.

Ann. Rev. Biochem. **50**: 879-910 (1981)

Gellert, M., Mizuuchi, K., O'Dea, M.H., Itoh, T. & Tomizawa, J. Novobiocin and coumermycin inhibit DNA supercoiling catalyzed by DNA gyrase.

Proc. Natl. Acad. Sci. USA **74**: 4772-4778 (1977)

Gibson, F.P., Leach, D.R.F. & Lloyd, R.G. Identification of *sbcD* mutations as cosuppressors of *recBC* that allow propagation of DNA palindromes in *Esherichia coli* K-12.

J. Bacteriol. **174**: 1222-1228 (1992)

Gilson, E., Bachellier, S., Perrin, S. et al. Palindromic unit highly repetitive DNA sequences exhibit species specificity within *Enterobacteriaceae*

Res. Micro. **141**: 1103-1116 (1990)

Glickman, B.W. & Ripley, L.S. Structural intermediates of deletion mutagenesis: a role for palindromic DNA.

Proc. Natl. Acad. Sci. USA **81**: 512-516 (1984)

Hirota, Y., Yasuda, S., Yamada, M., Nishimura, A., Sugimoto, K., Sugisaki, H., Oka, A. & Takanami, M. Structural and functional properties of the *Escherichia coli* origin of DNA replication.

Cold Spring Harbor Symp. Quant. Biol. **43**: 129-138 (1979)

Hulton, C. Seirafi, A., Hinton, J.C.D., Sidebottom, J.M., Waddell, L., Pavitt, G.D., Owen-hughes, T., Spassky, A., Buc, H. & Higgins, C.F. Histone-like protein HI (HNS), DNA supercoiling and gene expression in bacteria.

Cell **63**: 631-642 (1990)

Inselberg, J. Formation of deletion mutations in recombination-deficient mutants of *Escherichia coli.*

J. Bacteriol. **94**: 1266

Ikeda, H., Moriya, K. & Matsumoto, T. In vitro study of illegitimate recombination: involvement of DNA gyrase

Cold Spring Harbor Symp. Quant. Biol. **45**: 399-408 (1981)

Ikeda, H., Aoki, K. & Naito, A. Illegitimate recombination mediated in vitro by DNA gyrase of *Escherichia coli:* structure of recombinant molecules.

Proc. Natl. Acad. Sci. USA **79**: 3724-3728 (1982)

Ikeda, H., Kawasaki, I. & Gellert, M. Mechanism of illegitimate recombination: common sites for cleavage mediated by *Escherichia coli* DNA gyrase.

Mol. Gen. Genetics **196**: 546-549 (1984)

Ikeda, H. & Shiozaki, M. Nonhomologous recombination mediated by *Escherichia coli* DNA gyrase: possible involvement of DNA replication.

Cold Spring Harbor Symp. Quant. Biol. **49**: 401-409 (1984)

Ikeda, H. DNA topoisomerase mediated illegitmate recombination.

In: DNA Topology and its Biological Effects. (1990)

Jones, I.M., Primrose, S.B. & Ehrlich, S.D. Recombination between short direct repeats in a *RecA* host.

Mol. Ge. Genetics **188**: 486-488 (1982)

Joyce, C.M., Fujii, D.M., Laks, H.S., Hughes, C.M. & Grindley, N.D.F. Genetic mapping and DNA sequence analysis of mutations in the *polA* gene of *Escherichia coli.*

J. Mol. Biol. **186**: 283-293 (1985)

Kaiser, A. & Wu, R. Structure and function of DNA cohesive ends.

Cold Spring Harbor Symp. Quant. Biol. **33**: 729 (1968)

Karu, A.E., Sakaki, Y., Echols, H. & Linn, S. The gamma protein specified by bacterophage lambda. Structure and inhibitory activity for the recBC enzyme of *Escherichia coli.*

J. Biol. Chem. **250**: 7377-7387 (1975)

Kazic, T. & Berg, D.E. Context effects in the formation of deletions in *Escherichia coli.*

Genetics **126**: 17-24 (1990)

King, S.R., Krolewski, M.A., Marvo, S.L., Lipson, P.J., Pogue-Geile, K.L., Chung, J.H. & Jaskunas, S.R. Nucleotide sequence analysis of *in vivo* recombinants between bacteriophage lambda DNA and pBR322.

Mol. Gen. Genet. **186**: 548-557 (1982)

Kleckner, N. DNA sequence analysis of Tn10 insertions: origin and role of 9 bp flanking repetitions during Tn10 translocation.

Cell **16**: 711-720 (1979)

Kleckner, N. Transposon Tn10.

In: Mobile DNA pp.227-268 (1989) (American Society for Microbiology, Washington, DC)

Eds. Berg, D.E & Howe, M.M.

Kowalczykowski, S.C., Dixon, D.A., Eggleston, A.K., Lauder, S.D. & Rehrauer, W.M. Biochemistry of homologous recombination in *Escherichia coli.*

Microbiology Reviews **58**: 401-465 (1994)

Kulharni, S.K. & Stahl, F.W. Interaction between the *sbcC* gene of *Escherichia coli* and the *gam* gene of phage Lambda.

Genetics **123**: 249-253 (1989)

Kushner, S.R., Nagaishi, H., Templin, A. & Clark, A.J. Genetic recombination in *Escherichia coli:* the role of exonuclease I.

Proc. Natl. Acad. Sci. USA. **68**: 824-827 (1971)

Kuzminov, A. Collapse and repair of replication forks in *Escherichia coli.*

Mol. Micro. **16**: 373-384 (1995)

LaDuca, R.J., Fay, P.J., Chuang, C., McHenry, C.S. & Bambara, R.A. Site-specific pausing of deoxyribonucleic acid synthesis catalyzed by four forms of *Escherichia coli* DNA polymerase III.

Biochemistry **22**: 5177-5188 (1983)

Leach, D. & Lindsay, J. *In vivo* loss of supercoiled DNA carrying a palindromic sequence.

Mol. Gen. Genetics **204**: 322-327 (1986)

Leach, D., Lindsay, J. & Okley, E. Genome interactions which influence DNA palindrome mediated instability and inviability in *Escherichia coli.*

J. Cell. Sci. Suppl. **7**: 33-40 (1987)

Leach, D.R.F., & Stahl, F.W. Viability of λ phages carrying a perfect palindrome in the absence of recombination nucleases.

Nature **305**: 448-451 (1983)

Lejeune, P. & Danchin, A. Mutations in the *bglY* gene increase the frequency of spontaneous deletions in *Escherichia coli* K-12.

Proc. Natl. Acad. Sci. USA. **87**: 360-383 (1990)

Lloyd, R.G. & Buckman, C. Identification and genetic analysis of *sbcC* mutation in commonly used *recBC sbcB* strains of *Escherichia coli.*

J. Bacteriology **164**: 836-844 (1985)

Lockshon, D. & Morris, D.R. Sites of reaction of *Escherichia coli* DNA gyrase on pBR322 *in vivo* as revealed by oxolinic acid-induced plasmid linearization.

J. Mol. Biol. **181**: 63-74 (1985)

Lovett, S.T. & Kolodner, R.D. Identification and purification of a single-stranded-DNA-specific exonuclease encoded by the *recJ* gene of *Escherichia coli.*

Proc. Natl. Acad. Sci. USA **86**: 2627-2631 (1989)

Lu, A. L., Welsh, S., Clark, S., Su, S.S. & Modrich, P. Repair of DNA basepair mismatches in extracts of *Escherichia coli.*

Cold Spring Harbor Symp. Quant. Biol. **49**: 589-596 (1984)

Lundblad, V. & Kleckner, N. Mutants of *Escherichia coli* K12 which affect excision.

In: Molecular and Cellular Mechanisms of Mutatgenesis. p.245-258 (1982) (Academic Press, Inc., New York)

Eds. Lemontt, J.F. & Generoso, W.M.

Lundblad, V. & Kleckner, N. Mismatch repair mutations of*Escherichia coli* K12 enhance transposon excision.

Genetics **109**: 3-19 (1985)

Lundblad, V., Taylor, A.F., Smith, G.R. & Kleckner, N. Unusual alleles of *recB* and *recC* stimulate excision of inverted repeat transposons Tn10 and Tn5.

Proc. Natl. Acad. Sci. USA **81**: 824-828 (1984)

Maenhault-Michel, G. & Shapiro, J.A The role of starvation and selective substrates in the emergence of *araB-lacZ* fusion clones.

EMBO J. **13:** 5229-5239 (1994)

Mandel, M. & Higa, A. Calcium-dependent bacteriophage DNA infection.

J. Mol. Biol. **53**: 159-162 (1970)

Marvo. S.L., King, S.R. & Jaskunas, S.R Role of short regions of homology in intermolecular illegitimate recombination events.

Proc. Natl. Acad. Sci. USA **80**: 2452-2456 (1983)

Maxam, A.M. & Gilbert, W. A new method of sequencing DNA.

Proc. Natl. Acad. Sci. USA **74**: 560-564 (1977)

Mazin, A.V., Kuzminov, A.V., Dianov, G.L. & Salganik, R.I. Mechanisms of deletion formation in *Escherichia coli* plasmids. II. Deletions mediated by short direct repeats.

Mol. Gen. Genet. **228**: 209-214 (1991)

Mazin, A.V., Timchenko, T.V., Saparbaev, M.K. & Mazina, O.M. Dimerization of plasmid DNA accelerates selection for antibiotic resistance.

Mol. Micro. **20**: 101-108 (1996)

Meuth, M. Illegitmate recombination in mammalian cells.

In: Mobile DNA p833-860 (1989) (American Society for Microbiology, Washington, DC)

Eds. Berg, D.E. & Howe, M.M.

M'Graw, B.R. & Marinus, M.G. Isolation and characterization of *Dam*$^+$ revertants and suppressor mutations that modify secondary phenotypes of *dam-3* starins of *Escherichia coli.*

Mol. Gen. Genet. **178**: 309-315 (1980)

Mittler, J.E. & Lenski, R.E. Causes of mutation and Mu excision.

Nature **345**: 213 (1990)

Mizuuchi, M. & Mizuuchi, K. Target site selection in transposition phage Mu.

Cold Spring Harbor Symp. Quant. Biol. **58**: 515-523 (1993)

Naito, A., Naito, S. & Ikeda, H. Homology is not required for recombination mediated by DAN gyrase of *Escherichia coli.*

Mol. Gen. Genet. **193**: 238-243 (1984)

Philips, G.J., Prashner, D.C. & Kushner, S.R. Physiological and chemical characterisation of cloned *sbcB* and *xonA* mutations from *Escherichia coli* K12

J. Bacteriol. **170**: 2089-2094 (1988)

Pruss, G.J. & Drlica, K.

Cell **56**: 521-523 (1989)

Rebollo, J-E., Francois, V. & Louarn, J-M. Detection and possible role of two large nondivisible zones on the *Escherichia coli* chromosome.

Proc. Natl. Acad. Sci. USA **85**: 9391-9395 (1988)

Saing, K.M., Orii, H., Tanaka, Y., Yanagisawa, K. Miura, A. & Ikeda, H. Formation of deletion in *Escherichia coli* between direct repeats located in a long inverted repeats of cellular slime mold: participation of DNA gyrase.

Mol. Gen. Genet. **214**: 1624-1628 (1988)

Sanger, F., Nicklen, S. & Coulson, A.R. DNA sequencing with chain-terminating inhibitors.

Proc. Natl. Acad. Sci. USA **74**: 5463-5467 (1977)

Schaffermann, A., Flashner, Y., Hertman, I., Olami, Y. & Cohen, S. Molecular aspects of genetic instability of an artificial 68 bp perfect palindrome in *Escherichia coli.*

Mol. Gen. Genet. **208**: 294-300 (1987)

Scheurerman, R., Tam, S., Burgers, P.M.J, Lu, C. & Echols, H. Identification of the ε-subunit of *Eschericiha coli* DNA polymerase III holoenzyme as the *dnaQ* gene product: a fidelity subunit for DNA replication.

Proc. Natl. Acad. Sci. USA **80**: 7085-7089 (1983)

Schofield, M.A, Agbunag, R., Michaels, M.L. & Miller, J.H. Cloning and sequencing of *Escherichia coli mutR* shows its identity to *topB*, encoding top isomerase III.

J. Bacteriol. **174**: 5168-5170 (1992)

Shapiro, J.A. Observations on the formation of clones containing *araB-lacZ* cistron fusions.

Mol. Gen. Genet. **194**: 79-90 (1984)

Shapiro, J.A. A role for the Clp protease in inactivating Mu-mediated DNA rearrangements.

J. Bacteriol. **175**: 2625-2631 (1993)

Shapiro, J.A. & Leach, D.R.F. Action of a transposable element in coding sequence fusions.

Genetics **126**: 293-299 (1990)

Sharples, G.J. & Lloyd, R.G. A novel repeated DNA sequence located in the intergenic regions of bacterial chromosome.

Nucl. Acids. Res. **18**: 6503-6508 (1990)

Shen, P & Huang, H.V. Homologous recombination in *Escherichia coli*: genetic homology abd genetic length.

Genetics **112**: 441-457 (1986)

Shurvinton, C.E., Stahl, M.M. & Stahl, F.W Large palindromes in the lambda phage genome are preserved in a *rec*$^+$ host by inhibiting lambda replication.

Proc. Natl. Acad. Sci. USA **84**: 1624-1628 (1987)

Sinden, R.R & Pettijohn, D.D. Cruciform transitions in DNA.

J. Biol. Chem. **259**: 6593-6600 (1984)

Skarstad, K. & Boye, E. Perturbed chromosomal replication in *recA* mutants of *Escherichia coli.*

J. Bacteriol. **170**: 2549-2554 (1988)

Skarstad, K. & Boye, E. Degradation of individual chromosomes in *recA* mutants of *Escherichia coli.*

J. Bacteriol. **175**: 5505-5509 (1993)

Skarstad, K. Boye, E. & Steen, H.B. Timing of initiation of chromosomal replication in individual *Escherichia coli* cells.

EMBO J. **5**: 1711-1717 (1986)

Spudich, J.A., Horn, V. & Yanofsky, C. On the production of deletions in the chromosome of *Escherichia coli.*

J. Mol. Biol. **53**: 49-67 (1970)

Streisinger, G.S., Okada, Y., Emrich, J., Newton, J, Tsugita, A., Terzaghi, E. & Inouye, M. Frameshift mutations and the genetic code.

Cold Sring Harbor Symp. Quant. Biol. **31**: 77-84 (1967)

Summers, D.K., Benton, C.W.H. & Withers, H.L. Multicopy plasmid instability: the dimer catastrophe hypothesis.

Mol. Micro. **8**: 1031-1038 (1993)

Summers, D.F. & Sherratt, D.J. Multimerization of high copy number plasmids causes instability: ColE1 encodes a determinant essential for plasmid monomerization and stability.

Cell **36**: 1097-1103 (1984)

Syvanen, M., Hopkins, J.D., Griffin IV, T.J., Liang, T-Y., Ippen-Ihler, K. & Kolodner, R. Stimulation of precise excision and recombination bu conjugal proficient F' plasmids.

Mol. Gen. Genet. **203**: 1-7 (1986)

Trinh, T.Q. & Sinden, R.R. Preferential DNA secondary structure mutagenesis in the lagging strand of replication in *Escherichia coli.*

Nature **352**: 544-547 (1991)

Trinh, T.Q. & Sinden, R.R. The influence of primary and secondary DNA structure in deletion and duplication between direct repeats in *Escherichia coli.*

Genetics **134**: 409-422 (1993)

Ukita, T. & Ikeda, H. Role of *recJ* gene product in UV-induced illegitimate recombination at the hotspot.

J. Bacteriol. **178**: 2362-2367 (1996)

Vilette, D., Uzest, M., Ehrlich, S.D. & Michel, B. DNA transcription and repressor binding affect deletion formation in *Escherichia coli* plasmids.

EMBO J. **11**: 3629-3634 (1992)

Vogelstein, B. & Gillespie, D. Preparative and analytical purification of DNA from agarose.

Proc. Natl. Acad. Sci. USA **76**: 615-619 (1979)

Wang, J.C. DNA topoisomerases: why so many?

J. Biol. Chem. **266**: 6659-6662 (1991)

Warren, G.J. & Green, R.L. Comparison of physical and genetic properties of palindromic DNA sequences.

J. Bacteriol. **161**: 1103-1111 (1985)

Weaver, D.T & DePamphilis, M.L. The role of palindromic and non-palindromic sequences in arresting DNA synthesis *in vitro* and *in vivo*.

J. Mol. Biol. **180**: 961-986 (1984)

Weisberg, R.A. & Adaya, S. Illegitimate recombination in bacteria and bacteriophage.

Ann. Rev. Genet. **11**: 451-473 (1977)

Weisberg, R.A. & Landy, A. Site-specific recombination in phage lambda.

I: Lambda II p.211-250 (1983) (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY)

Eds. Hendrix, R.W., Roberts, J.W., Stahl, F.W. & Weisberg, R.A.

West, S.C The processing of recombination intermediates: Mechanistic insights from studies of bacterial proteins.

Cell **76** : 9-15 (1994)

Weston-Hafer, K & Berg, D.E. Palindromy and the location of deletion endpoints in *Escheichia coli.*

Genetics **121**: 651-658 (1989)

Weston-Hafer, K. & Berg, D.E. Limits to the role of palindromy in deletion formation.

J. Bacteriol. **127**: 315-318 (1991)

Whoriskey, S.K., Ngheim, V-H., Leong, P-M., Mason, J-M. & Miller, J.H. Gene amplification in *E. coli*: DNA sequences at the junctures of amplified gene fusions.

Genes Dev. **1**: 227-237 (1987)

Whoriskey, S.K., Schofield, M.A. & Miller, J.H. Isolation and characterization of *Escherichia coli* mutants with altered rates of deletion formation.

Genetics **127**: 21-30 (1991)

Young, Y. & Ames, G.F-L. The family of repetitive extragenic palindromic sequences:Interaction with DNA gyrase and histonelike protein HU.

In: The Bacterial Chromosome p. 211-225 (1990) (American Society for Microbiology, Washington, DC.)

Eds. Drlica, K. & Riley, M.

Yamaguchi, H., Yamashita, H., Shimizu, H. & Ikeda, H. A hotspot of spontaneous and UV-induced illegitimate recombination during formation of λ*bio* transducing phage.

Mol. Gen. Genet. **248**: 637-643 (1995)